

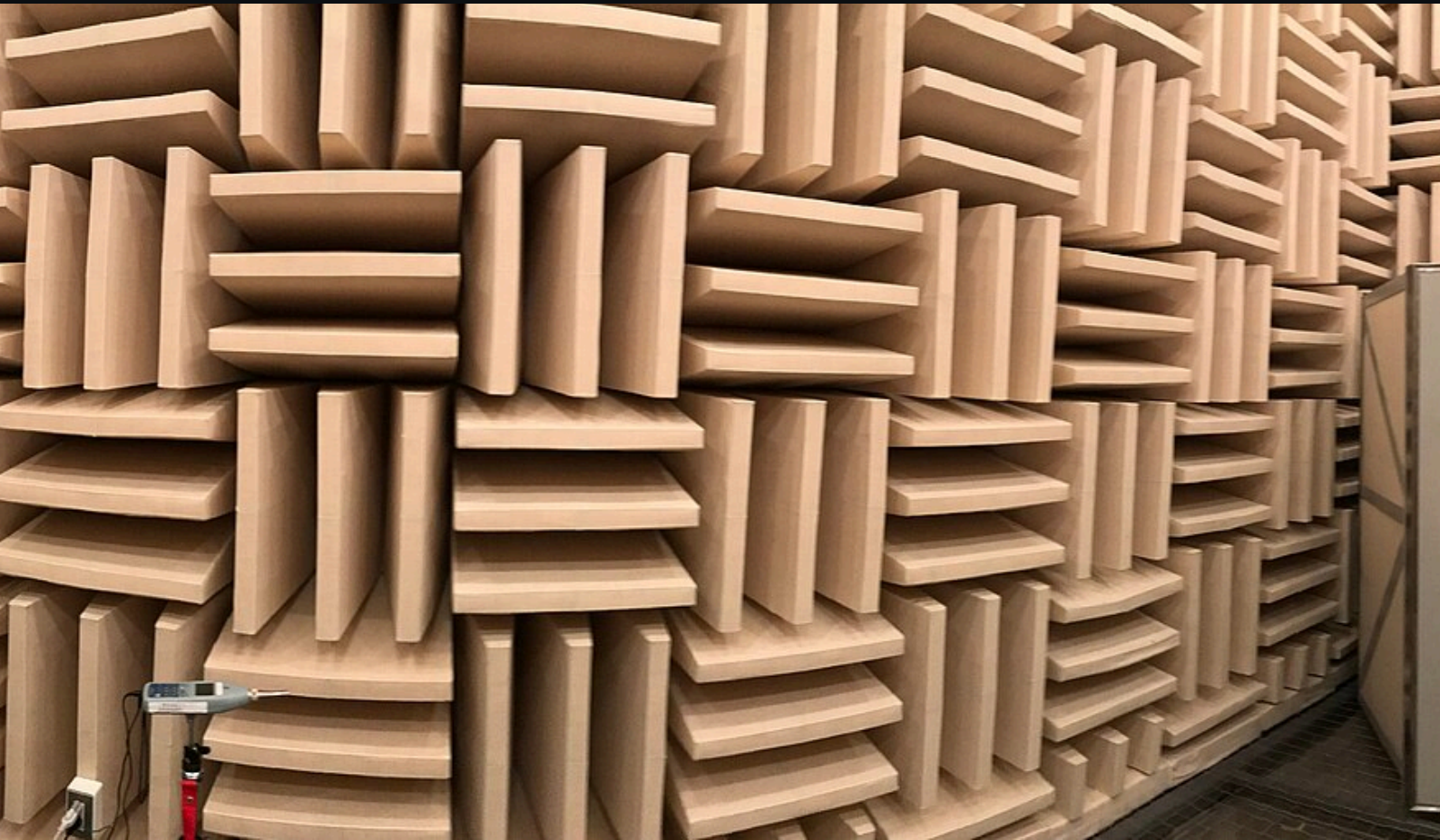


Photo: Anechoic chamber · Togabi · CC BY-SA 4.0 · Wikimedia Commons

OSINT · TECHNICAL ASSESSMENT

FROM THE LAB TO THE **FIELD**

Published accuracy figures for acoustic drone detection routinely sit at 97–99%. Deployed systems meet different microphones, different drones, wind, distance and a world full of things that hum. A 2026 Fraunhofer study measured how much performance survives the move from familiar to unfamiliar conditions, and what claws it back. This report sets out where the lab-to-field gap comes from, what fixes it, and how a buyer can test for it.

REPORT DETAILS

<div>PUBLISHED</div> <div>October 2026</div>	<div>SUBJECT</div> <div>The gap between laboratory and field performance in acoustic drone detection</div>	<div>SOURCE BASE</div> <div>Fraunhofer IDMT study (2026), AIP Advances review (2025), payload-detection and street-sound studies, outdoor microphone tests</div>
<div>METHOD</div> <div>Compare in-domain, out-of-domain, distance and false-alarm results from published evaluations</div>	<div>VERDICT</div> <div>Ask for out-of-domain, distance-binned results at a fixed false-alarm rate; headline accuracy alone means little</div>	<div>CONFIDENCE</div> <div>Moderate</div>
<div>AUTHOR</div> <div>Ghulam Mujtaba, Ph.D.</div>		

BOTTOM LINE

An acoustic drone detector is only as good as the conditions it was tested in, and most published results test it in the conditions it was trained in. Change the microphone, the site, the drone or the distance and accuracy falls, sometimes sharply. A May 2026 study from Germany's Fraunhofer Institute for Digital Media Technology put numbers on the drop. At a fixed 1% false-alarm rate, a detector caught 86% of drone sound on recordings like its training data, 70% on other drones, microphones and sites, and 17% on the hardest set. Two changes recovered most of the loss: pretraining the model on a large library of everyday sounds before teaching it drones, and deliberately roughening the training audio to resemble field conditions. The lesson for anyone buying or fielding acoustic detection is practical. Ask how a system performs on recordings it has never heard, at a stated false-alarm rate and at stated distances. A single accuracy figure answers none of those.

92%

DRONE SOUND CAUGHT ON FAMILIAR RECORDINGS

◆ PRETRAINED MODEL, 1% FALSE ALARMS

17→66%

ON THE HARDEST UNFAMILIAR SET

◆ TRAINED FROM SCRATCH VS PRETRAINED AND AUGMENTED

~36

FALSE 1-SECOND ALERTS AN HOUR AT A 1% RATE

◆ IF EVERY SEGMENT RAISED AN ALARM

100 m

HIGHLY RELIABLE RANGE IN ONE FIELD DATASET

◆ PARTLY EFFECTIVE TO ~150 M

KEY QUESTIONS

Why do acoustic drone detectors perform worse in the field than in tests?

Because most are trained and tested on recordings from the same microphones, sites and drones. In a 2026 Fraunhofer study, a detector that caught 86% of drone sound on familiar recordings caught 70% on recordings from other drones, microphones and sites, and 17% on the hardest set, at the same false-alarm setting. Pretraining on general sounds and field-style augmentation raised those to 92%, 83% and 66%.

How far away can an acoustic drone detector hear a drone?

In the Fraunhofer study, at a 1% false-alarm rate, detection was 98–99% within 50 m, 81–86% at 50–100 m, 52–68% at 100–150 m and near zero beyond 200 m. Longer listening windows extend reach at the cost of slower alerts.

What causes false alarms in acoustic drone detection?

Steady mechanical or buzzing sounds. A detector tuned to 1% false alarms on its home site flagged washing machines, chainsaws and insects as drones about 14–17% of the time on a public collection of everyday sounds.

THE RULER — THE WHOLE ARGUMENT

Where the lab and the field part company

Each way a deployment differs from a training set has a cost, and most have a known mitigation.

WHAT CHANGES — IN THE FIELD	WHY IT HURTS	WHAT HELPS
MICROPHONE DIFFERENT HARDWARE	Every microphone colours sound differently; models learn the colour	Simulating many microphone responses in training
SITE A NEW SOUNDSCAPE	Traffic, birds and machinery differ by place; models learn the place	Mixing varied background noise in training; testing on unseen sites
DRONE NEW MODELS, LOADS, SPEEDS	Rotor pitch shifts with size, payload and throttle	Pitch-shifting training audio; more drone types
DISTANCE TENS TO HUNDREDS OF METRES	The signal weakens and blurs into the background	Longer listening windows; denser sensor spacing
LOOK-ALIKES THINGS THAT HUM NEW	Washing machines, chainsaws and insects share the rotor's texture	Event logic over time; confirmation by a second sensor
RECORDING CHAIN COMPRESSION, CODECS	Web-sourced and compressed audio loses detail	Training on varied and degraded audio
EVERYTHING AT ONCE A LIVE FIELD TRIAL	Microphone, site, drone, distance and noise all differ from the lab together	Trials against live drones outdoors, as DroSonic ran at DIRT Days 26.04
● THE FAILURE THAT MATTERS MOST TO OPERATORS: FALSE ALARMS MITIGATIONS FROM THE STUDIES CITED BELOW		

FIELD 01 — THE BENCHMARK TRAP

Why 99% means less than it sounds

Near-perfect accuracy is common in the acoustic detection literature. Reviewing work on detecting whether a drone carries a payload, researchers at Université de Sherbrooke noted studies reporting around 99% accuracy, but with data from only one

or two drone models recorded at up to 15 metres; one study that measured distance found accuracy falling to 70% at 100 metres, and wind noise above 54 dB significantly hurting predictions. [Marcoux & Grondin](#) A December 2025 review of the whole field in *AIP Advances* makes the same point about the data underneath: most public drone-sound datasets cover common consumer multicopters, record within tens of metres, and follow no shared standard for equipment or labelling, which makes fair comparison between models difficult. [AIP Advances](#)

The pattern is not unique to drones. A 2024 study of street-sound classifiers built into city streetlights reported 85.4% accuracy on a standard benchmark, then found in real deployments that the model worked well for most vehicles but had significant limitations for pedestrians and bicycles. The authors named the quality and quantity of training data as the main limitation. [Discover Applied Sciences](#)

**A detector that has only heard its own microphone, its own field and its own drone
has learned those as much as it has learned drones.**

FIELD 02 — THE GAP, MEASURED

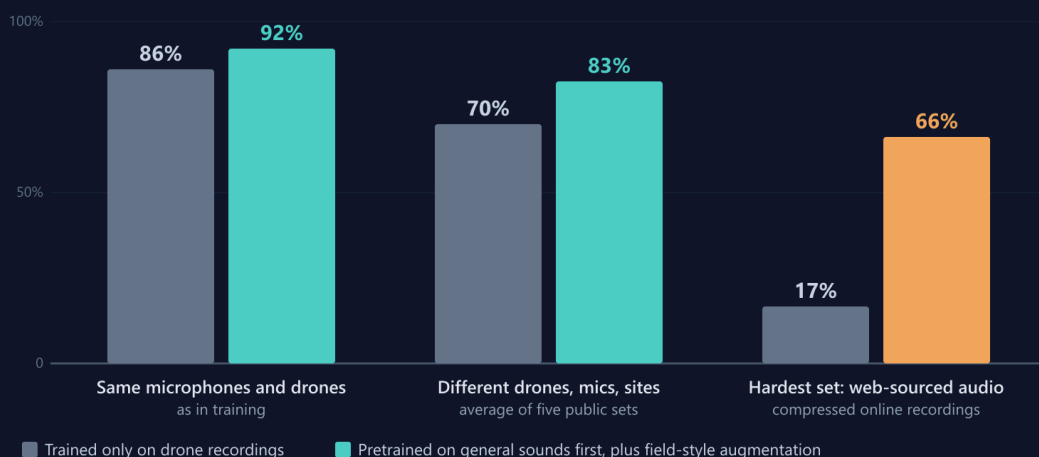
Familiar, unfamiliar and hardest

The Fraunhofer team trained a compact neural-network detector on 71 hours of drone recordings and 380 hours of background from their own eight-microphone array, plus public datasets, then tested it three ways. Crucially, they held out whole recording sessions, so the test never shared a day and place with training. They tuned every model to a fixed false-alarm rate on background audio and then measured how much drone sound it caught. [Fraunhofer IDMT](#)

Trained only on drone data, the detector caught 86% of drone segments on familiar recordings and 70% on public recordings made with other drones, microphones and sites. Pretraining it first on AudioSet, a large general collection of everyday sounds, lifted those to 92% and 78%; adding the full set of field-style augmentations brought the unfamiliar average to 83%. The hardest set, audio collected from the web, went from 17% to 66%. [Fraunhofer IDMT](#)

The same detector, three kinds of test

Share of drone sound detected while false alarms are held at 1% of background audio.



Source: Reuter, Ohlenbusch & Rollwage, Fraunhofer IDMT, arXiv:2605.31329 (May 2026), Table 3 and Fig. 1. Means of 10 runs.
Diagram: DroSonic OSINT.

osint.drosonic.com

ONE DETECTOR, THREE TESTS — THE DROP IS IN THE UNFAMILIAR

Pretraining on general sounds and field-style augmentation recover most of the loss, but not all of it.

DIAGRAM: DROSONIC OSINT · ORIGINAL · FROM REUTER, OHLENBUSCH & ROLLWAGE (FRAUNHOFER IDMT),
ARXIV:2605.31329

The authors attribute the remaining gaps to differences in drone types, recording hardware and post-processing, and suspect compression artefacts in web-sourced audio that their augmentations do not yet capture. [Fraunhofer IDMT](#)

FIELD 03 — DISTANCE

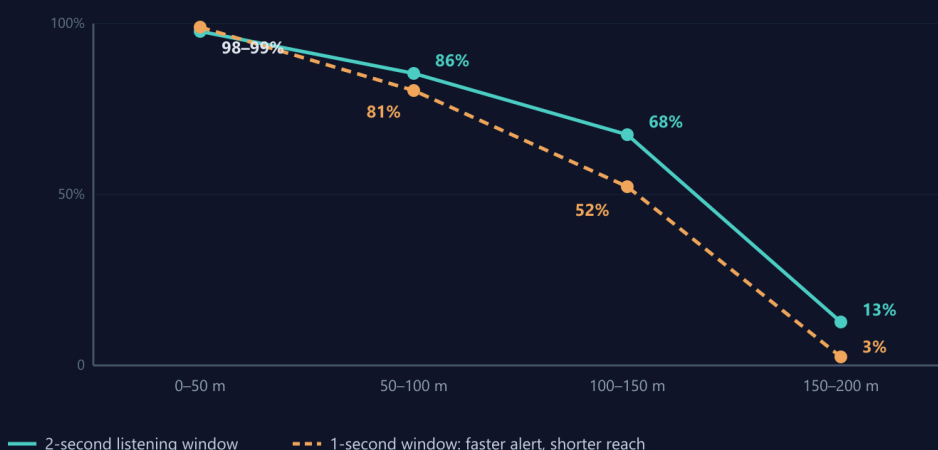
Where field performance leaks away

Range is the gap that no training trick removes. On real flights recorded in Berne, Germany, the Fraunhofer detector caught 98–99% of drone sound within 50 metres at a 1% false-alarm rate, about 81% at 50–100 metres, 52% at 100–150 metres and almost nothing beyond 200 metres. Listening for two seconds instead of one raised those middle figures to 86% and 68%, because a longer window gives a faint, wavering signal more chance to stand out. The cost is a slower alert. The authors call the detector highly reliable to 100 metres and partly effective to about 150. [Fraunhofer IDMT](#)

Denser spacing is the route DroSonic takes: many small networked nodes rather than one large array, so no drone has to be heard from far away. [DroSonic](#)

Range is where field performance leaks away

Share of drone sound detected at a 1% false-alarm rate, by distance from the microphone.



Source: Reuter, Ohlenbusch & Rollwage, Fraunhofer IDMT, arXiv:2605.31329 (May 2026), Sec. 5.3; real flights recorded in Berne, Germany, 2022.
Diagram: DroSonic OSINT. Near zero beyond 200 m at this false-alarm setting.

osint.drosonic.com

DETECTION BY DISTANCE — TIME BUYS RANGE

A longer listening window reaches further but alerts later. Beyond about 150 m, sensor spacing matters more than the model.

DIAGRAM: DROSONIC OSINT · ORIGINAL · FROM REUTER, OHLENBUSCH & ROLLWAGE (FRAUNHOFER IDMT),
ARXIV:2605.31329

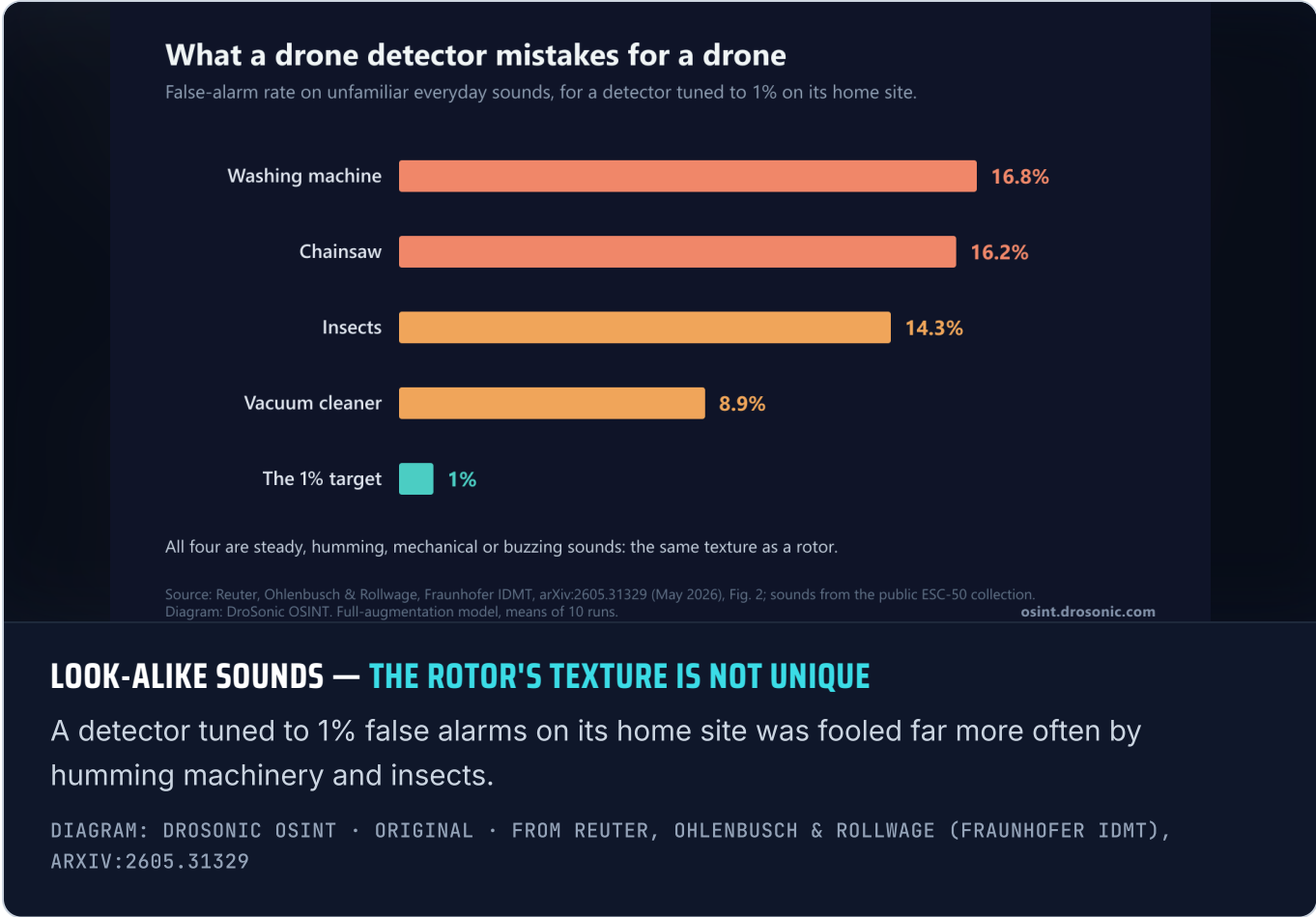
These ranges come from one microphone array, one site and a few camera drones (DJI Phantom, DJI Mavic and Align models). [Fraunhofer IDMT](#) ⤴ Louder drones carry further, and quieter sites help, but the shape of the curve is the point: a single acoustic node covers a neighbourhood, not a region. Our explainer [What a Microphone Can Hear](#) sets out the physics of why.

FIELD 04 — FALSE ALARMS

The world is full of things that hum

An operator's tolerance for false alarms is far lower than a benchmark's. The Fraunhofer authors spell out what a 1% false-alarm rate means at the scale of individual one-second decisions: roughly 36 false segments an hour, which would be too many if each raised an alert. They argue that detections need merging over time and filtering by further logic, or confirmation by another sensor, with acoustics serving as a sensitive early cue. [Fraunhofer IDMT](#) ⤴

What fools a drone detector is telling. Tested on a public collection of 2,000 everyday sound clips, the model most often mistook washing machines, chainsaws and insects for drones, at 14–17% each, followed by vacuum cleaners at 9%. Every one is a steady mechanical or buzzing texture, like a rotor. [Fraunhofer IDMT](#)



Radar has its own version of this problem, with birds; our report [Is It a Bird?](#) covers it. The cure is the same in both cases: a second, independent signature before anyone acts.

FIELD 05 — WEATHER

The anechoic chamber has no wind

Laboratory recordings are often made in quiet rooms, or close to the drone on still days. Outdoors, wind blowing across the microphone is one of the strongest noise sources, and it sits in the same low frequencies as rotor tones. Physical protection

helps a great deal: one outdoor study of drone-detection microphones found that a windscreen extended detection range in strong wind by roughly 31–131%, depending on the wind's direction. [PubMed Central](#)↗



A MICROPHONE WINDSHIELD — **CHEAP, LOW-TECH, DECISIVE**

Fur and foam slow the air at the capsule. For outdoor acoustic sensing, mechanical design matters as much as the model.

PHOTO: NICOLAI EQUIPMENT · CC BY-SA 3.0 · [WIKIMEDIA COMMONS](#)

FIELD 06 — WHAT CLOSES THE GAP

Training for the world, testing like a buyer

The Fraunhofer results rank the fixes. **Pretraining** on a broad library of sounds was the single largest gain, on familiar and unfamiliar data alike. **Augmentation** came second and mattered most on mismatched data: simulating different microphone responses gave the biggest single jump, then mixing in background noise, masking parts of the spectrum, and pitch-shifting drone sound to mimic other rotor speeds. [Fraunhofer IDMT](#)↗

Augmentation is not free: the most aggressively augmented model raised false alarms on unfamiliar everyday sounds slightly, from about 2% to 2.5% at the 1% setting. [Fraunhofer IDMT](#)↗

For a buyer, the questions follow directly:

1. **Was the test data recorded on different days, places and microphones from the training data?** If not, the result is a lab result.
2. **What was the false-alarm rate when the detection rate was measured?** A detection rate without one is meaningless.
3. **How does detection fall with distance?** Ask for results in distance bands, not an average.
4. **Which sounds cause false alarms, and what happens next?** Look for event logic over time and a second sensor before action.
5. **Which drones were tested?** Size, payload and propulsion change the sound.

CONFIDENCE

How sure we are, and what would change it

High confidence that acoustic detectors lose accuracy on unfamiliar microphones, sites and drones, and with distance. The pattern appears across independent studies, reviews and an unrelated street-sound deployment.

Moderate confidence in the specific figures, which come largely from one study with one array, a few camera drones and segment-level scoring. Event-level performance in an operational system, with time-merging and sensor fusion, will differ.

What would change it: published, independent field trials of acoustic detectors at fixed false-alarm rates on unseen sites and drone types, which remain rare in the open literature.

DROSONIC — OUR APPROACH

Field data over benchmark scores

Every row of the table above applies to DroSonic's own sensors, which is why its development moved from recordings to open ground early. Detection runs as neural networks on low-cost edge hardware inside each node, so the model, the microphone and the site are tested together rather than one at a time. [DroSonic](#)

In September 2026 a prototype was tested against live drones at DIRT Days 26.04, with wind, vehicles and an electronic environment built to resemble a contested battlefield; DroSonic reports that it heard drones other sensors missed. [DroSonic](#) That is an account, not a published benchmark. The test this report recommends, detection by distance at a fixed false-alarm rate on unseen sites, is the standard any buyer should hold acoustic detection to, DroSonic's included.

CLOSE – WHAT IT MEANS

Test where it will be used

None of this argues against acoustic detection. It argues for honesty about it. The same study that shows the gap shows that most of it can be closed with known methods, and that what remains is mainly range and look-alike sounds, both of which a system design can absorb: denser nodes, longer listening where speed allows, and a second sensor to confirm before anyone acts.

At [DroSonic](#) that is why we test outdoors, against live drones and in noise, not only on recordings. In September 2026 a DroSonic prototype was tested against live drones at the DIRT Days 26.04 field exercise. [DroSonic](#)



THE OTHER END OF THE SCALE — A FIELD TEST, NOT A QUIET ROOM

DroSonic's test position at DIRT Days 26.04: open ground, wind, vehicles and live drones. Results from days like this, not from the lab, are what a buyer should ask for.

PHOTO: DROSONIC · [FIELD-TEST WRITE-UP](#)

ABOUT DROSONIC

[DroSonic](#) builds passive acoustic drone detection: networked sensors that listen for drones instead of their radio signals, using acoustic AI and edge computing. DroSonic OSINT is our open-source research into the threat that technology is built to meet.

THE RECORD SO FAR

OCT 2026	Joined the SOF RACER consortium, which connects USSOCOM with commercial technology companies
SEPT 2026	Prototype tested against live drones at DIRT Days 26.04, a field exercise supported by CMI2 with U.S. Central Command and U.S. Army DEVCOM
SEPT 2026	Technical presentation at NDIA's Emerging Technologies for Defense conference on detecting fibre-optic and radio-silent drones
SEPT 2026	Filed U.S. provisional patent 64/148,522 on passive acoustic detection, classification, localisation and defeat cueing

JUL 2026

Pitched to the U.S. Space Force at Space Systems Command's Reverse Industry Engagement Day, Colorado Springs

[How DroSonic detection works →](#) · [Where it is used →](#) · [Talk to the team →](#)

THIS REPORT ONLINE

<https://osint.drosonic.com/assessments/acoustic-detection-lab-vs-field.html>

Cite as: Mujtaba, G. (2026, October 30). *From the Lab to the Field: Why Acoustic Drone Detectors Fail in the Field*. DroSonic OSINT. <https://osint.drosonic.com/assessments/acoustic-detection-lab-vs-field.html>

Reusing this report. Text and original diagrams © DroSonic. You may quote this report and share its link with credit to DroSonic OSINT; to republish it in full, reuse our diagrams or use it commercially, [contact us](#) first. Photos and third-party figures keep the licences shown in their credits.

DroSonic OSINT: open-source assessments on drones and counter-UAS from DroSonic, passive acoustic drone detection. drosonic.com · sudo@drosonic.com



ABOUT THE AUTHOR

Ghulam Mujtaba, Ph.D.

FOUNDER AND CEO, DROSONIC

Ghulam Mujtaba, Ph.D., is the founder and CEO of [DroSonic](#) and leads its work on passive acoustic drone detection. A computer engineer with 13 years of experience building machine-learning systems for edge hardware, he holds a Ph.D. in computer engineering from Gachon University in South Korea, was an NSF postdoctoral fellow at West Virginia University in the United States, and is an Assistant Professor of AI and Data Science at Regis University in Denver, Colorado. An IEEE Senior Member, he has published more than 20 papers on deep learning, computer vision and multimedia for edge devices.

[PROFILE ↗](#) · [LINKEDIN ↗](#) · [ORCID ↗](#)

ALL SOURCES (7)

1. [Marcoux & Grondin](#) — Prior payload studies ~99% at short range, one or two drone models; 70% at 100 m; wind above 54 dB
2. [AIP Advances](#) — Kang et al. 2025 review: datasets within tens of metres, consumer multirotors, no unified standards
3. [Discover Applied Sciences](#) — Goulão et al. 2024: 85.4% on ESC-50; real-world limits for 'person' and 'bicycle'; data the main limitation

4. Fraunhofer IDMT — 71 h drone / 380 h background; session-disjoint split; thresholds calibrated to 1% and 5% FPR
5. DroSonic — Passive acoustic detection: neural networks on low-cost edge hardware, no emissions
6. PubMed Central — Windscreen extended range in strong wind by ~31-131%
7. DroSonic — DIRT Days 26.04 field test, September 2026