Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Predicting missed alerts in detection dogs

Jörg Schultz

Joerg.Schultz@Tier-Wohl-Team.de

Tier Wohl Team GbR

Liza Rothkoff

Texas Tech University

Edgar O. Aviles-Rosa

Texas Tech University

Nathaniel J. Hall

Texas Tech University

Michele N. Maughan

Precise Systems Inc.

Article

Keywords:

Posted Date: June 16th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-9811291/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Predicting missed alerts in detection dogs

Jörg Schultz[1,*], Liza Rothkoff[2], Edgar O. Aviles-Rosa[2], Nathaniel J. Hall[2], and Michele N. Maughan[3]

[1]Tier Wohl Team GbR, Rödelsee, Germany
[2]Department of Animal and Food Science, Texas Tech University, Lubbock, TX, United States
[3]Precise Systems Inc., United States
[*]Joerg.Schultz@Tier-Wohl-Team.de

ABSTRACT

Detection dogs are widely used in tasks ranging from explosives detection to medical diagnostics. Yet even well-trained dogs occasionally fail to indicate target odors. Such missed alerts can have severe operational consequences. Here we show that missed alerts are preceded by measurable changes in the dog's behavioral and physiological state. Using a controlled olfacto-treadmill system with precisely timed odor presentation, we analyzed movement dynamics and physiological signals during the seconds preceding each trial. Pre-odor movement patterns predicted subsequent missed indications with above-chance accuracy (AUC = 0.81). Incorporating physiological signals increased recall for missed alerts from 77% to 85%. These signals emerged seconds before odor exposure, indicating that detection failures reflect underlying variation in the dog's physiological and attentional state. When aggregated across trials, model predictions provided a continuous estimate of this evolving state. Together, these findings suggest that detection performance reflects a dynamic biological process and provide a foundation for anticipatory monitoring of working dog readiness.

Introduction

Detection dogs are widely used to identify target odors in tasks ranging from explosives detection to wildlife conservation and medical diagnostics[1–5]. Through training, dogs learn to make an indication response in the presence of a specific volatile chemical signature, allowing them to convert trace-level odor information into an observable indication for a human handler. In operational settings, these indications often guide critical decisions about the presence of hazardous substances, missing persons, or disease-related biomarkers. Nonetheless, like all biological systems, performance can fluctuate based on the vigilance, attentional and physical state of the detector[6–9].

Operational errors in detection work occur in two principal forms: (i) false alerts in the absence of a target odor and (ii) missed alerts when a target is present but not indicated. While both types of error have operational implications, missed alerts are often of particular concern in high-stakes environments, where failure to detect explosives, hazardous substances, or disease-related odors may carry serious consequences. Importantly, similar to human operators, dogs can show waning detection capabilities based on their psychological state, such as vigilance and attention[6,7,10] and physiological state related to ongoing physical work demands[11,12].

One potentially unexplored explanation for variability in working dog performance lies in the physiological demands of detection work. Sustained locomotion, thermoregulatory load, and repeated odor discrimination tasks impose measurable physiological strain. Experimental studies have demonstrated that exercise intensity and duration can impair olfactory detection limits, particularly at low odor concentrations, even before overt behavioral signs of fatigue become apparent[11]. Alterations in autonomic regulation as reflected in heart rate dynamics and heart rate variability[13–15], as well as increases in core body temperature[16–20], may precede observable changes in search behavior. Together, these findings suggest that fluctuations in detection performance may arise from gradual shifts in physiological state. In humans, variability in task performance has been linked to changes in underlying physiological state, particularly in sustained attention paradigms[21], and such fluctuations can manifest in subtle changes in movement patterns, as internal state modulates motor behavior including gait dynamics[22].

In operational settings, the primary safeguard against performance decline is the handler's ongoing assessment of the dog. Handlers monitor posture, gait, search intensity, response latency, and overall engagement, integrating these cues with knowledge of the individual dog's training history and working style. Decisions to pause, rotate, or rest a dog are therefore grounded in expert judgment. This observational approach is indispensable and often effective. However, subtle physiological changes that precede overt behavioral alterations may not be readily visible, particularly in complex or high-pressure environments where attention is divided among navigation, communication, and safety considerations. Consequently, early stages of performance decline may remain difficult to detect.

Experimental research shows that workload and exercise influence autonomic regulation, thermoregulatory load, and movement dynamics—factors closely linked to olfactory performance[11,12,23,24]. These findings suggest that changes in quantifiable biological processes could be used as predictors of performance. Thus, identifying subtle behavioral changes associated with these physiological processes is critical for detecting early performance decline. The presence of these behaviors could be used by handlers to prevent odor misses before they happen.

To evaluate such signals, the temporal window relative to odor exposure must be precisely defined. In unconstrained search environments, the moment of first odor encounter is typically unknown, making it difficult to isolate behavior occurring before stimulus contact. Experimental investigation of pre-response signals therefore requires precise control over odor onset. To address this, we employed a controlled olfacto-treadmill system that enables computer-regulated odor delivery with precise temporal onset[11]. By combining an automated olfactometer with treadmill locomotion, dogs can be exercised under standardized conditions while receiving odor stimuli at known time points. The fixed lateral position of the dog further permits synchronized recording of video and physiological signals with minimal variability across trials and individuals. This controlled framework provides the temporal precision required to isolate and quantify pre-odor movement and physiological signals.

With this experimental precision in place, our goal was to test whether measurable variations in movement and physiology before odor exposure encode information about impending performance failure. If declining performance reflects biological processes that unfold prior to overt errors, then miss risk should be detectable from pre-odor signals alone. Demonstrating such a relationship would indicate that detection performance can be inferred from an anticipatory state rather than from post hoc evaluation.

Results

Pre-odor Movement Predicts Missed Alerts

To assess whether impending detection failures are reflected in signals preceding odor contact, we analyzed data from a controlled olfacto-treadmill paradigm in which dogs performed repeated odor detection trials while moving at a fixed gait[25] (Fig. 1). Dogs were tested both at walking (4 km/h) and trotting speed (8 km/h). Because performance decline was more pronounced under higher workload conditions, all analyses presented here were restricted to target-present trials performed at 8 km/h. We then analyzed the 4.5 s immediately preceding each target odor presentation. Behavioral outcomes (hit vs. miss) were used only as training labels.

Movement was quantified using markerless pose estimation[26] to extract 29 anatomically defined landmarks per frame (Fig. 1). Each trial yielded a sequence of 135 frames (30 Hz), capturing joint-level posture and coordination over the pre-odor window. For each landmark, two-dimensional coordinates and a prediction confidence were recorded, resulting in a time series of 29 landmarks with three features each per frame. Coordinates were normalized relative to the withers to remove positional bias on the treadmill. This representation converts raw video into a time-resolved skeletal description of the dog's posture and movement, enabling analysis of subtle shifts in head position, axial alignment, limb coordination, and tail dynamics.

Because anatomical landmarks form an articulated body structure, joints cannot be treated as independent variables. Instead, landmarks were represented as nodes in a graph reflecting skeletal connectivity, allowing the model to capture coordinated posture changes across the body. This graph representation was processed using a Graph Attention Network (GAT), which enables the model to learn relationships between connected joints[27]. Movement patterns also evolve continuously over time rather than occurring in isolated frames. To capture these temporal dynamics, the sequence of pose graphs was processed by a Temporal Convolutional Network (TCN), integrating movement information across the entire 4.5 s window[28]. Together, this spatiotemporal architecture (Fig. 2) allows the model to detect both brief micro-adjustments and slower drift in posture preceding odor exposure.

The pose-only model was trained on augmented, class-balanced pre-odor sequences, with equal representation of hit and miss trials. On the independent test set (216 trials: 144 non-indications, 72 indications), pre-odor movement predicted missed alerts with above-chance accuracy (AUC = 0.81; accuracy = 73.6%). Validation performance and training dynamics are provided in Supplementary Table S2 and Fig. S2.

Performance was strongest for missed alerts (non-indications), with a precision of 0.82 and a recall of 0.77. This indicates that most predicted misses were true misses and that the majority of actual misses were correctly identified. In contrast, performance for indication trials was lower (precision = 0.59, recall = 0.67), reflecting a conservative bias toward predicting elevated-risk states.

Because the model had no access to odor cues or trial outcomes at inference time, predictive performance must derive from movement structure expressed before stimulus contact. These findings demonstrate that subtle variations in posture and coordination during the pre-odor phase contain predictive information about subsequent detection failure.

Physiological Signals Enhance Prediction of Missed Alerts

Although movement thus contains predictive information, it reflects only the externally observable expression of behavior. Physiological measurements provide a complementary view of the internal state of the animal, including autonomic regulation and thermoregulatory load, both of which have been associated with performance variability in working dogs. We therefore examined whether incorporating physiological signals recorded during the same 4.5 s pre-odor window improves predictive performance.

To this end, we extended the pose-based model by adding three physiological inputs: heart rate (HR), inter-beat interval (IBI) statistics capturing short-term heart rate variability, and core body temperature. These signals were recorded concurrently with video and aligned to the identical pre-stimulus window. Unlike the pose data, which form a frame-wise time series across the 4.5 s window, physiological features were represented as summary measures over the interval and processed through a separate network branch. The resulting physiological representation was then fused with the movement-based embedding prior to classification (Fig. 2), allowing joint modeling of external movement dynamics and internal physiological state.

Incorporating physiological signals increased sensitivity to missed alerts while maintaining overall discrimination. On the independent test set, the multimodal model achieved an accuracy of 76.9% and an AUC of 0.80, comparable to the video-only model (AUC = 0.81). Validation performance and training dynamics are provided in Supplementary Table S2 and Fig. S3.

The primary effect of physiological integration was an increase in recall for missed alerts, rising from 0.77 in the video-only model to 0.85 in the multimodal model. Precision for this class remained stable (0.82 vs. 0.81), indicating that the improved sensitivity did not come at the cost of increased false positives. Correspondingly, the F1-score for missed alerts increased from 0.80 to 0.83. Performance for indication trials remained lower (precision = 0.67, recall = 0.61), consistent with a conservative bias toward predicting elevated-risk states. Overall, physiological signals enhanced detection of trials culminating in missed alerts without materially affecting specificity.

These findings show that short segments of autonomic and thermoregulatory data recorded before odor presentation provide complementary predictive information beyond movement alone. Incorporating physiological signals increases sensitivity to impending detection failures without materially affecting precision.

Behavioral and Physiological Features Underlying Prediction

Axial posture dominates movement-based prediction

While the preceding analyses demonstrate that pre-odor movement contains predictive information about missed alerts, they do not reveal which aspects of behavior change before a detection failure occurs. Identifying these behavioral signals is essential for interpreting the results in biological terms and for understanding what observable cues might indicate declining readiness in a working dog.

To identify which body regions carry this predictive information, we applied a gradient-based saliency analysis to the video-only model[29]. For each test trial, we quantified how small perturbations in individual joint coordinates affected the predicted probability of a missed alert and averaged these effects across frames and trials to obtain a mean saliency score for each landmark (Fig. 3A). Saliency values were rescaled for readability, with higher values indicating joints whose variation exerted a stronger influence on model output. A dynamic visualization of this saliency analysis is provided in Supplementary Video S2.

Landmarks along the axial body line clearly dominated the saliency ranking. The top of the head (mean saliency: 0.212), the second cervical vertebra (axis; 0.132), and the withers (0.128) exerted the strongest influence on model output, indicating sensitivity to subtle changes in head and neck posture. Although the withers were used as the coordinate origin during preprocessing, the graph attention architecture propagates information from neighboring landmarks; saliency at this node therefore reflects the importance of relative axial posture encoded in the surrounding joints. The tip of the tail also ranked prominently (0.100), and the tip of the nose showed a moderate contribution (0.051). In contrast, most distal limb joints and thoracolumbar landmarks showed substantially lower influence (generally <0.034). One anatomically occluded landmark (right ischial tuberosity; 0.187) also exhibited elevated saliency. Because the dogs were filmed from the left side, this landmark was not directly visible and is therefore not interpreted as an observable behavioral feature. Its apparent importance may instead reflect unstable attribution under occlusion or indirect propagation of information from neighboring joints within the graph-based model. Accordingly, this landmark was interpreted with caution.

The concentration of predictive signal along the head–neck–trunk–tail axis is biologically plausible. Head posture directly reflects orientation toward the odor port and may encode subtle changes in engagement, attentional focus, or anticipatory sampling behavior during the pre-odor phase. Small shifts in head elevation or axial alignment may signal declining readiness before overt behavioral errors occur. Tail position and movement have been linked to arousal and emotional state in dogs[30–32]. That these regions emerged as dominant contributors without explicit anatomical weighting suggests that the model converged on behaviorally meaningful posture dynamics. Predictive information in pre-odor movement therefore appears to be concentrated in axial posture-related regions associated with engagement and arousal rather than distributed uniformly across the body.

Heart rate variability is the primary physiological driver

To assess the contribution of individual physiological inputs, which form a small set of distinct modalities, we conducted a systematic ablation analysis[33]. In contrast to the pose representation, where landmarks are structurally interdependent within a spatiotemporal graph, physiological signals can be meaningfully removed at the modality level, enabling direct evaluation of their effects. Each physiological modality was removed in turn while all other inputs were retained, and predictive performance was reassessed on the test set using AUC, classification accuracy, and log-loss for evaluation (Fig. 3B).

Removing IBI produced the largest degradation in predictive performance. The area under the ROC curve (AUC) decreased by 0.16 (95% CI: 0.08–0.23), accuracy dropped by 6 percentage points, and log-loss increased by 0.19 relative to the full multimodal model. Eliminating core temperature resulted in a more moderate performance decrease (ΔAUC = 0.06, 95% CI: 0.03–0.10; accuracy -4.1%; log-loss +0.20). In contrast, removing heart rate alone had minimal impact on discriminative performance: AUC and accuracy remained essentially unchanged, although log-loss increased (Δlog-loss = 0.49), indicating reduced probability calibration without a shift in the decision boundary.

This hierarchy is biologically informative. IBI captures short-term modulation of autonomic nervous system activity and is closely linked to arousal, stress regulation, and cognitive load in dogs[13,14]. Its dominance indicates that impending missed alerts are associated with alterations in autonomic regulation rather than simple physical exertion. Core temperature provides additional context related to physiological strain but contributes less strongly to discriminative prediction. Heart rate alone—often used as a workload proxy—adds comparatively little predictive value within this anticipatory window.

Together, these findings suggest that subtle shifts in autonomic balance precede detection failures and are detectable before odor contact.

Converging markers of declining readiness

Taken together, the behavioral and physiological analyses successfully predict impending misses. Predictive information in behavior is concentrated in head posture and tail dynamics—external markers of engagement and arousal—while predictive information in physiology is dominated by short-term heart rhythm variability, an indicator of autonomic regulation.

The convergence of these signals suggests that declining detection performance is not triggered by a single behavioral event but reflects a broader change in the dog's physiological and attentional state. Importantly, these markers emerge within the 4.5 s window preceding odor presentation, before the dog has encountered the stimulus.

By linking predictive performance to specific behavioral and physiological features, the framework moves beyond opaque classification toward a biologically interpretable description of early performance decline.

Aggregated Model Outputs Track Evolving Performance State

Single-trial predictions fluctuated and therefore did not directly capture gradual changes in performance within a session. We therefore examined whether trial-level model outputs could be aggregated to approximate an evolving performance state across consecutive trials. Model confidence scores were treated as probabilistic evidence and sequentially integrated across target-present trials using a Beta-based accumulation framework. This procedure yields a continuously updated estimate of the underlying miss probability, producing a smooth trajectory of inferred miss risk over time (Fig. 4A).

As an illustrative example, we analyzed a complete run from one dog comprising 50 target-present trials (Fig. 4B). For each trial, the model generated a confidence score based solely on the 4.5 s pre-odor window. These scores were aggregated sequentially to estimate the probability that the underlying miss rate exceeded predefined thresholds. This trajectory was compared to the observed miss frequency calculated over sliding windows of 10 trials. As the empirical miss frequency increased during the session, the aggregated model output increased in parallel. Because model predictions were generated without access to trial outcomes, this correspondence reflects information contained in pre-odor movement and physiological signals.

In this example, a neutral prior was used, corresponding to no strong initial assumption about the dog's miss rate. Alternative prior specifications produced qualitatively similar trajectories (Fig. S4).

These observations demonstrate that aggregating trial-level predictions can capture gradual changes in performance within a session. However, this analysis is based on a single illustrative example and does not constitute a systematic evaluation across dogs or conditions.

Discussion

A central question motivating this study was whether fluctuations in detection performance reflect an underlying biological state that becomes detectable before an error occurs. Using a controlled olfacto-treadmill system that allowed precise isolation of the seconds preceding odor exposure, we found that movement patterns alone contained sufficient information to predict subsequent missed alerts. Predictive sensitivity increased further when physiological signals were incorporated. Because all input data were recorded strictly before odor delivery and without access to behavioral outcomes, the predictive signal must arise from

variations in the dog's behavioral and physiological state rather than from reactions to odor processing. These findings indicate that missed alerts are preceded by measurable changes in engagement-related posture and autonomic regulation, supporting the view that detection performance reflects a dynamic biological process rather than isolated stochastic failures.

The predictive signal was not diffuse across the body but concentrated in anatomically and behaviorally meaningful regions, particularly cranial landmarks and the tip of the tail. Subtle changes in head posture may reflect altered attentional orientation or sampling readiness during the pre-odor phase, whereas tail dynamics are widely associated with arousal and motivational state in dogs[30–32]. The prominence of these regions suggests that missed alerts are preceded by subtle deviations in engagement-related posture rather than by gross locomotor disruption. Physiological measurements provided further insight. Short-term heart rhythm variability contributed more strongly than raw heart rate, indicating that regulatory dynamics within the autonomic nervous system contain informative signal within the pre-stimulus window. This pattern argues against a simple explanation based on physical fatigue alone and instead points to shifts in autonomic regulation that may accompany changes in cognitive load, arousal balance, or sustained task demand[13,14].

Importantly, these alterations emerged while the dogs continued trotting under constant workload, without visible disruption of gait or stability. This indicates that missed alerts are not preceded by overt behavioral breakdown but by more subtle perturbations in the biological systems supporting sustained detection work. Detection tasks—particularly at low odor concentrations—place continuous demands on attentional and regulatory processes. Under such conditions, even modest changes in physiological or cognitive state may influence the probability of a successful indication.

If detection performance fluctuates as a biological state, the question becomes how such changes should be interpreted during ongoing work. In practice, handlers are less concerned with isolated events than with the dog's current state. The relevant question is therefore: What is the likelihood that my dog is operating in a condition where missed alerts exceed an acceptable threshold? This is the level at which deployment decisions—continue searching or allow recovery—are typically made.

Interpreting model outputs at this level therefore requires moving beyond single-trial predictions toward an estimate of the dog's evolving performance state. In the present study, we aggregated consecutive trial-level predictions to estimate whether the dog's underlying miss rate was increasing beyond an acceptable threshold. By integrating model outputs across trials, this approach translates fluctuating single-trial predictions into a probabilistic assessment of performance state. Such an estimate can incorporate prior information, including the dog's historical performance or the expected difficulty of the task, and allows acceptable risk thresholds to be adjusted depending on operational consequences. Rather than prescribing a fixed decision rule, the framework provides a structured way to translate anticipatory biological signals into context-sensitive deployment guidance (Fig. 5).

The controlled conditions that made it possible to isolate signals also differ substantially from real-world search environments. All data were collected under laboratory conditions using the olfacto-treadmill[11]. This design enabled precise isolation of the pre-odor window and standardized measurement of movement and physiology, but it does not capture the complexity of real-world search environments. Field deployments involve variable terrain, handler interaction, environmental distractions, and heterogeneous target distributions that were not represented here.

A further practical constraint concerns the use of video-based pose estimation[26]. Continuous, unobstructed video capture is rarely feasible during operational searches. However, the present findings suggest that the predictive signal resides in the structure of movement rather than in the video modality itself. Wearable inertial sensors and accelerometers are already widely used in canine research and may provide a more field-compatible means of capturing analogous movement signatures[34–36]. Replicating the present analyses using wearable data represents a logical next step, although it will require adaptation of the modeling framework to different signal characteristics.

A related consideration concerns the temporal scope of the analyzed signals. The present study focused on short pre-odor intervals defined by the experimental structure of the task, in which odor presentations occurred at fixed time points. This design enabled precise isolation of anticipatory signals while minimizing contamination from preceding trial outcomes, which themselves can influence the dog's immediate behavioral and physiological state. As a result, the model operates on temporally localized snapshots of the dog's condition immediately prior to each odor encounter.

It is notable that predictive information was detectable within this brief window, indicating that measurable markers of declining readiness can emerge on a sub-trial timescale. At the same time, performance fluctuations are likely to evolve over longer periods spanning multiple trials. Future work may therefore examine how signals at different temporal scales—from short anticipatory intervals to longer-term trends—can be combined. Importantly, such approaches would complement rather than replace the present framework, in which aggregation of short-window predictions provides a mechanism for tracking gradual changes in performance state across a session.

The olfacto-treadmill nevertheless provides a valuable translational bridge. By allowing precise experimental control, it enables identification of candidate behavioral and physiological markers of declining readiness before testing them in progressively more complex environments. Future work should evaluate whether similar anticipatory patterns can be detected

across larger populations, diverse task types, and operationally realistic conditions.

In summary, we show that missed alerts are preceded by measurable shifts in posture and autonomic regulation. This indicates that declining readiness becomes detectable immediately prior to a detection error. Detection performance therefore appears to reflect an evolving biological state rather than a sequence of isolated successes and failures. While further validation will be required before operational deployment, these findings suggest that monitoring such signals may enable earlier and more informed decisions about whether a dog should continue working or be allowed to recover, with implications for both operational reliability and animal welfare.

Methods

Ethics Statement

The study was conducted at the Texas Tech University Canine Olfaction Research and Education (CORE) Lab and approved by the Texas Tech University Institutional Animal Care and Use Committee (IACUC) and the U.S. Army Animal Care and Use Review Office (ACURO). All methods were performed in accordance with the relevant guidelines and regulations.

Data Collection and Preprocessing

Subjects

Four dogs participated in this study (Table 1). All dogs were between 2 and 4 years of age, healthy, and actively trained for detection work. One dog (Charles) was the pet dog of a laboratory staff member and was brought to the facility daily for training and testing. The remaining three dogs (Adele, Boomer, and Iitooma) were procured for the project from Chiron K9 (Somerset, TX, USA) and were housed at the Texas Tech University Canine Olfaction Research and Education (CORE) Lab facility. One dog (Charles) had also participated in our prior treadmill study[11]. Dogs were selected based on their familiarity with treadmill walking and consistent participation across sessions. This information has previously been reported in a companion manuscript[20].

Experimental Setup and Data Acquisition

Olfacto-treadmill setup and task design. All recordings were obtained during a treadmill-based odor discrimination task using a custom-built olfacto-treadmill[11,20]. Dogs maintained a consistent gait—either walking (4 km/h) or trotting (8 km/h)—throughout the session, while odors were delivered via an automated olfactometer. All analyses in the present study were restricted to target-present trials performed at 8 km/h.

Trial structure. Each trial began with computer-controlled activation of an odor valve (target or diluent). Dogs were trained to insert their muzzles into an odor port, breaking an infrared beam. Responses were scored as hits if the beam was broken for a dog-specific minimum duration (0.4 s for Boomer and Charles; 0.8 s for Adele and Iitooma). A failure to respond within 10 s was scored as a miss. Correct responses were reinforced with food. Trials were initiated every 20 s, yielding fixed onset intervals and uniform session durations. Target presence was randomized at 50% frequency. The experimenter was blind to trial type; valve selection and feedback tones were fully automated. The target odor was 1-bromooctane (1-BO), diluted in mineral oil (v/v).

Session formats. Two session types contributed data: (i) *Fixed-pace sessions* included 100 trials at a constant gait (walk or trot), lasting 33 min 20 s. (ii) *Pace-transition sessions* included 140 trials (approximately 46 min) with a mid-session gait switch (walk $\rightarrow$ trot or trot $\rightarrow$ walk) after trial 70. Each dog completed four fixed-pace sessions (two walk, two trot) and eight pace-transition sessions (four of each transition type) on separate days, with at least 18 h between sessions. Exercise was voluntary; if a dog stepped off the treadmill, the session resumed once the dog re-engaged.

Video recording. Video was recorded from a fixed lateral camera positioned at mid-body height under consistent indoor lighting and background. Recordings were time-synchronized with the olfactometer and physiological event logs. For each included trial, the 4.5 s immediately preceding odor onset (defined by the olfactometer valve activation timestamp) were extracted for modeling. Details of pose estimation and feature extraction are provided in the following section.

Physiological measurements. Core body temperature was measured using an ingestible AniPill^{TM} capsule (BodyCap, France), sampled every minute. Heart rate (HR) and inter-beat intervals (IBI) were continuously recorded using a Polar® H10 chest strap and logged at 1 Hz. All physiological data were aligned to trial timing. Temperature values were temporally aligned to trial onset using the closest preceding measurement.

Subset used for modeling. To isolate predictors of missed indications under moderate workload, analyses were restricted to target-present trials at 8 km/h. Trials from both fixed-pace and pace-transition sessions were pooled for analysis, as all included trials were performed at the same gait (8 km/h) and extracted using an identical pre-odor window. For each included trial, video and physiological data were aligned to the 4.5 s pre-odor window, and behavioral labels (hit vs. miss) were assigned based on the infrared beam-break criteria described above.

Physiological Data Processing

Gastrointestinal temperature data were sampled at 1 min intervals and linearly interpolated to 1 s resolution. Heart rate (HR) was recorded continuously at 1 Hz. All physiological signals were time-aligned to olfactometer valve activation timestamps defining odor onset. For each included trial, physiological data were extracted from the 4.5-s interval immediately preceding odor onset. Only pre-stimulus data were used for feature extraction. Trials containing signal loss, sensor dropout, or incomplete physiological recordings within the pre-odor window were excluded from analysis.

Trial Labeling and Ground Truth

Each trial was labeled based on the dog's behavioral response during odor presentation. Only target-present trials were included in the modeling analysis; diluent control trials were excluded. A nose poke meeting the predefined infrared beam-break duration criterion was labeled as a *hit*, whereas failure to meet this criterion within the 10 s response window was labeled as a *miss*. Labels were derived directly from the experiment control software using infrared beam-break event logs. Behavioral labels were used exclusively for model training and evaluation; no outcome information was available during extraction of pre-odor movement or physiological features.

Data Preprocessing and Inclusion Criteria

Only target-present trials with complete and valid pose and physiological recordings were included in the analysis. For each retained trial, a 5 s interval immediately preceding odor onset (defined by the olfactometer valve activation timestamp) was initially extracted.

From this 5 s segment, a 4.5 s window was used for model input. For non-augmented samples, the final 4.5 s immediately preceding odor onset were selected. For augmented samples (see Supplementary Section S2), a randomly shifted 4.5 s subwindow within the original 5 s segment was extracted to reduce alignment bias across trials. All features were aligned to the selected 4.5 s window prior to preprocessing.

Pose data. The initial 5-s segment contained 150 frames sampled at 30 Hz. The final model input consisted of 135 consecutive frames (4.5 s). Pose annotations included 29 anatomical landmarks per frame (see Fig. 1), each represented by two pixel coordinates (x, y) and a likelihood value. To remove global positional bias, coordinates were translated such that the withers (joint 4) were located at the origin in each frame. This normalization reduced sensitivity to absolute body position on the treadmill and within the video frame, emphasizing relative joint movement.

Heart rate. Heart rate was sampled at 1 Hz. Values corresponding to the selected 4.5 s window were extracted from the original 5 s segment. Missing values due to brief signal loss were linearly interpolated.

Inter-beat interval (IBI). Inter-beat intervals within the selected 4.5 s window were used to compute seven summary statistics: (1) mean RR interval, (2) standard deviation, (3) minimum, (4) maximum, (5) root-mean-square of successive differences (RMSSD), (6) coefficient of variation, and (7) mean absolute successive difference. Trials containing fewer than two RR intervals within the window were excluded.

Core temperature. For each trial, the most recent temperature measurement at or preceding odor onset was selected.

All features (pose coordinates, heart rate, IBI statistics, and temperature) were z-scored per dog to reduce inter-individual differences in baseline physiology and body size.

Pose Estimation and Segmentation-Assisted Refinement

Movement was quantified using markerless pose estimation to obtain a time-resolved skeletal representation of each dog during treadmill locomotion. We trained a DeepLabCut-based model[26] to track 29 anatomical landmarks spanning the entire body, including cranial reference points, axial markers along the spine, forelimb and hindlimb joints, and tail landmarks (Fig. 1). The landmark set was designed to provide comprehensive anatomical coverage without imposing region-specific weighting or prior assumptions regarding which body segments might be predictive.

A total of 120 frames were manually annotated across dogs and sessions to capture a broad range of postures and movement phases. A single model was trained and subsequently applied across all individuals. For each frame (30 Hz), the network produced two-dimensional landmark coordinates and associated likelihood values.

To improve robustness under varying background conditions and partial occlusion, pose estimation was performed on segmentation-refined frames (details in Supplementary Section S1; validation in Supplementary Fig. S1 and Table S1). An initial pose prediction was used to define a region of interest centered on the dog's body. Within this region, a segmentation model was applied to isolate the dog from the background[37], and pose estimation was repeated on the masked image. The region of interest was updated dynamically across frames based on the preceding pose prediction, enabling stable tracking throughout each sequence.

All pose sequences used in downstream analyses were derived from these segmentation-refined predictions. A visual example of the segmentation-assisted refinement procedure, including landmark overlay and background removal, is provided in Supplementary Video S1.

Model Architecture and Training

Spatiotemporal Graph Network (ST-GAT-TCN)

Pose-only model. To model structured movement dynamics, we implemented a spatiotemporal graph neural network (ST-GAT-TCN)[38] (Fig. 2). Each input sample consisted of a sequence of 135 consecutive frames (4.5 s at 30 Hz), with 29 anatomical landmarks per frame. For each landmark, two-dimensional coordinates (x, y) and a confidence value were used, yielding an input tensor of shape $(135, 29, 3)$.

Spatial dependencies between joints were modeled using a fixed undirected graph with 29 nodes, corresponding to the anatomical landmarks. Edges reflected musculoskeletal connectivity, including spinal chains and limb–torso attachments. This graph structure was applied consistently across all time steps. Frame-wise spatial processing was performed using stacked graph attention layers, enabling adaptive weighting of neighboring joints while preserving anatomical structure.

Temporal dynamics were modeled using a multi-layer temporal convolutional network (TCN) operating along the time axis. Dilated one-dimensional convolutions were employed to capture movement patterns across multiple temporal scales within the 4.5 s pre-odor window. The resulting temporal representation was globally pooled and passed through fully connected layers, with a final sigmoid output producing the predicted probability of an indication.

Multimodal model with physiological input. To incorporate internal physiological state, the pose-based ST-GAT-TCN architecture was extended with a parallel physiological branch (Fig. 2). For each trial, heart rate, inter-beat interval (IBI) summary statistics, and core body temperature extracted from the 4.5-s pre-odor window were used as additional inputs.

Physiological features were processed through dedicated feedforward layers to generate a compact physiological embedding. This embedding was combined with the spatiotemporal pose representation via late fusion prior to the final classification layer. The fused representation was passed to the same sigmoid output head used in the pose-only model.

This modular design preserved the spatiotemporal backbone of the pose-based network while enabling joint modeling of behavioral movement dynamics and autonomic physiological state. Additional architectural and training details are provided in Supplementary Section S2.

Training and Evaluation

Model development was restricted to target-present trials with complete pose and physiological data available within the 4.5 s pre-odor window. Trials were first partitioned into a held-out test set and a development set. The test set was not used during model training, augmentation, or model selection and contained only original, unmodified samples.

Within the development set, data augmentation was applied to increase robustness and reduce sensitivity to fixed temporal alignment within the pre-odor window. For each trial, additional samples were generated by extracting randomly shifted 4.5-s subwindows from the original 5-s clip and applying small spatial perturbations to joint coordinates. Augmentation was performed separately within each class, and class balancing was achieved by matching the number of samples across indication and miss trials.

The resulting development data were then randomly partitioned at the sample level into training and validation subsets. Model selection was based on validation AUC. Each architecture was trained multiple times with different random initializations to ensure stability of performance estimates. Run-wise validation metrics are summarized in Supplementary Table S2.

Final model performance was reported exclusively on the independent held-out test set using accuracy, area under the receiver operating characteristic curve (AUC), precision, and recall.

Interpretability Analyses

Saliency analysis of pose features. To identify which anatomical landmarks most strongly influenced movement-based predictions, we analyzed the trained pose-only model using a gradient-based saliency approach[29]. For each trial in the held-out test set, the gradient of the model output with respect to each joint's input features (two-dimensional coordinates and detection confidence) was computed in evaluation mode without updating model weights. Absolute gradient magnitudes were averaged across frames within the 4.5 s pre-odor window and subsequently averaged across trials to obtain a global importance score per landmark.

Physiological sensor ablation. The contribution of physiological signals was assessed using a sensor ablation procedure[33]. For each physiological modality (heart rate, inter-beat interval features, and core body temperature), inputs from the corresponding branch were replaced with zeros while all other inputs were left unchanged. Using the same trained multimodal model, predictions were recomputed on the held-out test set and compared against the full-input condition using AUC, accuracy, and log-loss.

Bayesian Aggregation of Trial-Level Predictions

To approximate an evolving performance state across consecutive trials, we modeled the dog's underlying miss probability within a session as an unknown latent parameter $\theta \in [0, 1]$. Within a given session, θ was assumed constant but updated

sequentially as probabilistic evidence accumulated from trial-level model outputs. A Beta prior distribution, $\theta \sim \text{Beta}(\alpha_0, \beta_0)$, was specified to represent baseline uncertainty regarding miss probability at session onset. Unless otherwise stated, a weakly informative prior was used ($\alpha_0 = 1, \beta_0 = 1$), corresponding to a uniform distribution over $[0,1]$. Further mathematical details and prior-sensitivity analyses are provided in Supplementary Section S3.

For each target-present trial i, the trained model generated a predicted probability p_i of a missed alert based exclusively on the 4.5 s pre-odor window. These predicted probabilities were treated as fractional Bernoulli observations and used as soft evidence to update the Beta posterior sequentially according to:

$$\alpha_i = \alpha_{i-1} + p_i$$

$$\beta_i = \beta_{i-1} + (1 - p_i)$$

This update yields a closed-form posterior distribution $\theta \mid \text{data} \sim \text{Beta}(\alpha_i, \beta_i)$ after i trials. At each step, the posterior mean miss probability was computed as:

$$\mathbb{E}[\theta \mid \text{data}] = \frac{\alpha_i}{\alpha_i + \beta_i}$$

In addition, exceedance probabilities were calculated to quantify the likelihood that the underlying miss rate exceeded a predefined operational threshold τ. Specifically, $P(\theta > \tau \mid \text{data})$ was computed analytically from the cumulative Beta distribution. To facilitate comparison with observed behavioral performance, empirical miss frequency was calculated using a trailing sliding window of 10 consecutive target-present trials. Aggregated posterior trajectories were then compared qualitatively to the sliding-window miss rate within the same session.

References

1. Furton, K. G. & Myers, L. J. The scientific foundation and efficacy of the use of canines as chemical detectors for explosives1. *Talanta* **54**, 487–500, DOI: 10.1016/S0039-9140(00)00546-4 (2001).
2. Adams, G. J. & Johnson, K. G. Sleep, work, and the effects of shift work in drug detector dogs *Canis familiaris*. *Appl. Animal Behav. Sci.* **41**, 115–126, DOI: 10.1016/0168-1591(94)90056-6 (1994).
3. Jones, K. E., Dashfield, K., Downend, A. B. & Otto, C. M. Search-and-rescue dogs: An overview for veterinarians. *J. Am. Vet. Med. Assoc.* **225**, 854–860, DOI: 10.2460/javma.2004.225.854 (2004).
4. McKeague, B., Finlay, C. & Rooney, N. Conservation detection dogs: A critical review of efficacy and methodology. *Ecol. Evol.* **14**, e10866, DOI: 10.1002/ece3.10866 (2024).
5. Maughan, M. N. *et al.* The Use and Potential of Biomedical Detection Dogs During a Disease Outbreak. *Front. Medicine* **9**, 848090, DOI: 10.3389/fmed.2022.848090 (2022).
6. Pattyn, N., Neyt, X., Henderickx, D. & Soetens, E. Psychophysiological investigation of vigilance decrement: Boredom or cognitive fatigue? *Physiol. & Behav.* **93**, 369–378, DOI: 10.1016/j.physbeh.2007.09.016 (2008).
7. Aviles-Rosa, E. O., DeChant, M. T., Prada-Tiedemann, P. A. & Hall, N. J. A laboratory model of canine search vigilance decrement, I. *J. Exp. Analysis Behav.* **120**, 103–119, DOI: 10.1002/jeab.832 (2023).
8. Fisk, A. D. & Schneider, W. Control and Automatic Processing During Tasks Requiring Sustained Attention: A New Approach to Vigilance. *Hum. Factors* **23**, 737–750, DOI: 10.1177/001872088102300610 (1981).
9. Adams, J. A. Vigilance in the detection of low-intensity visual stimuli. *J. Exp. Psychol.* **52**, 204–208, DOI: 10.1037/h0048556 (1956).
10. Porritt, F. *et al.* Performance decline by search dogs in repetitive tasks, and mitigation strategies. *Appl. Animal Behav. Sci.* **166**, 112–122, DOI: 10.1016/j.applanim.2015.02.013 (2015).
11. Aviles-Rosa, E. *et al.* A canine model to evaluate the effect of exercise intensity and duration on olfactory detection limits: The running nose. *Front. Allergy* **5**, DOI: 10.3389/falgy.2024.1367669 (2024).
12. Gazit, I. & Terkel, J. Explosives detection by sniffer dogs following strenuous physical activity. *Appl. Animal Behav. Sci.* **81**, 149–161, DOI: 10.1016/S0168-1591(02)00274-5 (2003).
13. Katayama, M. *et al.* Heart rate variability predicts the emotional state in dogs. *Behav. Process.* **128**, 108–112, DOI: 10.1016/j.beproc.2016.04.015 (2016).

14. Craig, L., Meyers-Manor, J. E., Anders, K., Sütterlin, S. & Miller, H. The relationship between heart rate variability and canine aggression. *Appl. Animal Behav. Sci.* **188**, 59–67, DOI: 10.1016/j.applanim.2016.12.015 (2017).

15. Dickinson, S. & Feuerbacher, E. N. Frustration and its impact on search and rescue canines. *Front. Vet. Sci.* **12**, DOI: 10.3389/fvets.2025.1546412 (2025).

16. Angle, T. C. & Gillette, R. L. Telemetric measurement of body core temperature in exercising unconditioned Labrador retrievers. *Can. J. Vet. Res. = Revue Can. De Recherche Veterinaire* **75**, 157–159 (2011).

17. Baker, J. *et al.* Body Temperature Responses During Phases of Work in Human Remains Detection Dogs Undergoing a Simulated Deployment. *Animals: an open access journal from MDPI* **10**, 673, DOI: 10.3390/ani10040673 (2020).

18. Carter, A. J. & Hall, E. J. Investigating factors affecting the body temperature of dogs competing in cross country (canicross) races in the UK. *J. Therm. Biol.* **72**, 33–38, DOI: 10.1016/j.jtherbio.2017.12.006 (2018).

19. O'Brien, C., Karis, A. J., Tharion, W. J., Sullivan, H. M. & Hoyt, R. W. Core temperature responses of military working dogs during training activities and exercise walks. *U.S. Army Med. Dep. J.* 71–78 (2017).

20. Rothkoff, L. *et al.* Exercise intensity and duration impacts on gastrointestinal temperature and odor detection performance of dogs, DOI: 10.2139/ssrn.5481748 (2025).

21. van den Berg, R. *et al.* A common mechanism underlies changes of mind about decisions and confidence. *eLife* **5**, e12192, DOI: 10.7554/eLife.12192 (2016).

22. Yogev-Seligmann, G., Hausdorff, J. M. & Giladi, N. The role of executive function and attention in gait. *Mov. Disord. Off. J. Mov. Disord. Soc.* **23**, 329–342; quiz 472, DOI: 10.1002/mds.21720 (2008).

23. Altom, E. K., Davenport, G. M., Myers, L. J. & Cummins, K. A. Effect of dietary fat source and exercise on odorant-detecting ability of canine athletes. *Res. Vet. Sci.* **75**, 149–155, DOI: 10.1016/S0034-5288(03)00071-7 (2003).

24. Angle, C. T. *et al.* The effects of exercise and diet on olfactory capability in detection dogs. *J. Nutr. Sci.* **3**, e44, DOI: 10.1017/jns.2014.35 (2014).

25. Rothkoff, L. *et al.* A preliminary study on the impacts of exercise intensity and duration on gastrointestinal temperature and odor detection performance of dogs. *Appl. Animal Behav. Sci.* **299**, 106970, DOI: 10.1016/j.applanim.2026.106970 (2026).

26. Mathis, A. *et al.* DeepLabCut: Markerless pose estimation of user-defined body parts with deep learning. *Nat. Neurosci.* **21**, 1281–1289, DOI: 10.1038/s41593-018-0209-y (2018).

27. Veličković, P. *et al.* Graph Attention Networks, DOI: 10.48550/arXiv.1710.10903 (2018).

28. Lea, C., Vidal, R., Reiter, A. & Hager, G. D. Temporal Convolutional Networks: A Unified Approach to Action Segmentation. In Hua, G. & Jégou, H. (eds.) *Computer Vision – ECCV 2016 Workshops*, 47–54, DOI: 10.1007/978-3-319-49409-8_7 (Springer International Publishing, Cham, 2016).

29. Simonyan, K., Vedaldi, A. & Zisserman, A. Deep Inside Convolutional Networks: Visualising Image Classification Models and Saliency Maps. In Bengio, Y. & LeCun, Y. (eds.) *2nd International Conference on Learning Representations, ICLR 2014, Banff, AB, Canada, April 14-16, 2014, Workshop Track Proceedings* (2014).

30. Quaranta, A., Siniscalchi, M. & Vallortigara, G. Asymmetric tail-wagging responses by dogs to different emotive stimuli. *Curr. biology: CB* **17**, R199–201, DOI: 10.1016/j.cub.2007.02.008 (2007).

31. Beerda, B., Schilder, M. B. H., van Hooff, J. A. R. A. M., de Vries, H. W. & Mol, J. A. Behavioural, saliva cortisol and heart rate responses to different types of stimuli in dogs. *Appl. Animal Behav. Sci.* **58**, 365–381, DOI: 10.1016/S0168-1591(97)00145-7 (1998).

32. Leonetti, S., Cimarelli, G., Hersh, T. A. & Ravignani, A. Why do dogs wag their tails? *Biol. Lett.* **20**, 20230407, DOI: 10.1098/rsbl.2023.0407 (2024).

33. Li, J., Monroe, W. & Jurafsky, D. Understanding Neural Networks through Representation Erasure. *CoRR* **abs/1612.08220** (2016).

34. Reinstein, R. d. S. *et al.* Analysis of dog movement using a single accelerometer in different body positions—a new approach. *Front. Vet. Sci.* **12**, DOI: 10.3389/fvets.2025.1551341 (2025).

35. Chambers, R. D. *et al.* Deep Learning Classification of Canine Behavior Using a Single Collar-Mounted Accelerometer: Real-World Validation. *Animals* **11**, 1549, DOI: 10.3390/ani11061549 (2021).

36. Bolton, S., Cave, N., Cogger, N. & Colborne, G. R. Use of a Collar-Mounted Triaxial Accelerometer to Predict Speed and Gait in Dogs. *Animals* **11**, 1262, DOI: 10.3390/ani11051262 (2021).

37. Lugaresi, C. *et al.* MediaPipe: A Framework for Building Perception Pipelines, DOI: 10.48550/arXiv.1906.08172 (2019).

38. Grattarola, D. & Alippi, C. Graph Neural Networks in TensorFlow and Keras with Spektral [Application Notes]. *Comp. Intell. Mag.* **16**, 99–106, DOI: 10.1109/MCI.2020.3039072 (2021).

Funding

Funding was provided by the U.S. Army via the Chemical Biological Advanced Materials and Manufacturing Science (CBAMMS) Program (PE 0601102A Project VR9) at the Combat Capabilities Development Command Chemical Biological Center (DEVCOM CBC).

Author contributions statement

Conceptualization: Jörg Schultz, Edgar Aviles-Rosa, Nathaniel J. Hall, Michele N. Maughan

Data curation: Jörg Schultz

Formal analysis: Jörg Schultz

Funding acquisition: Michele N. Maughan

Investigation: Liza Rothkoff, Edgar Aviles-Rosa, Nathaniel J. Hall

Methodology: Jörg Schultz, Liza Rothkoff, Edgar Aviles-Rosa, Nathaniel J. Hall, Michele N. Maughan

Project administration: Michele N. Maughan

Software: Jörg Schultz

Visualization: Jörg Schultz

Writing - original draft: Jörg Schultz

Writing - review & editing: Jörg Schultz, Liza Rothkoff, Edgar Aviles-Rosa, Nathaniel J. Hall, Michele N. Maughan

Additional information

Data and Code Availability

All code, trained models, and example data required to reproduce the analyses are available at Zenodo: https://doi.org/10.5281/zenodo.17197670 (version v1.0.0).

Competing interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Figures

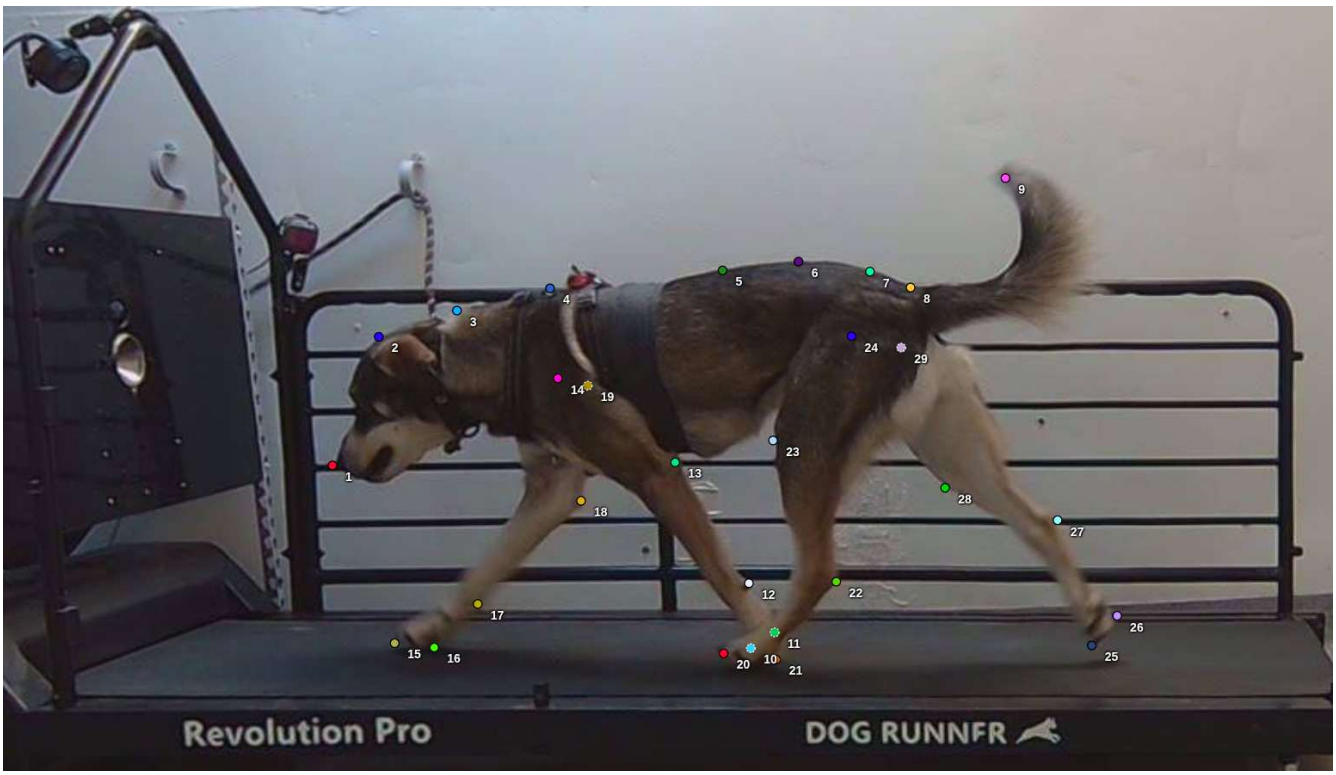

Figure 1. Markerless pose representation of a detection dog during treadmill-based odor detection. Full-body anatomical landmarks annotated on a detection dog during a controlled olfacto-treadmill task with automated odor delivery. Twenty-nine landmarks were defined across the head and neck (1–3), axial body (4–7), tail (8–9), forelimbs (10–19), and hindlimbs (20–29). A complete list of landmark descriptions is provided in Supplementary Table 1.

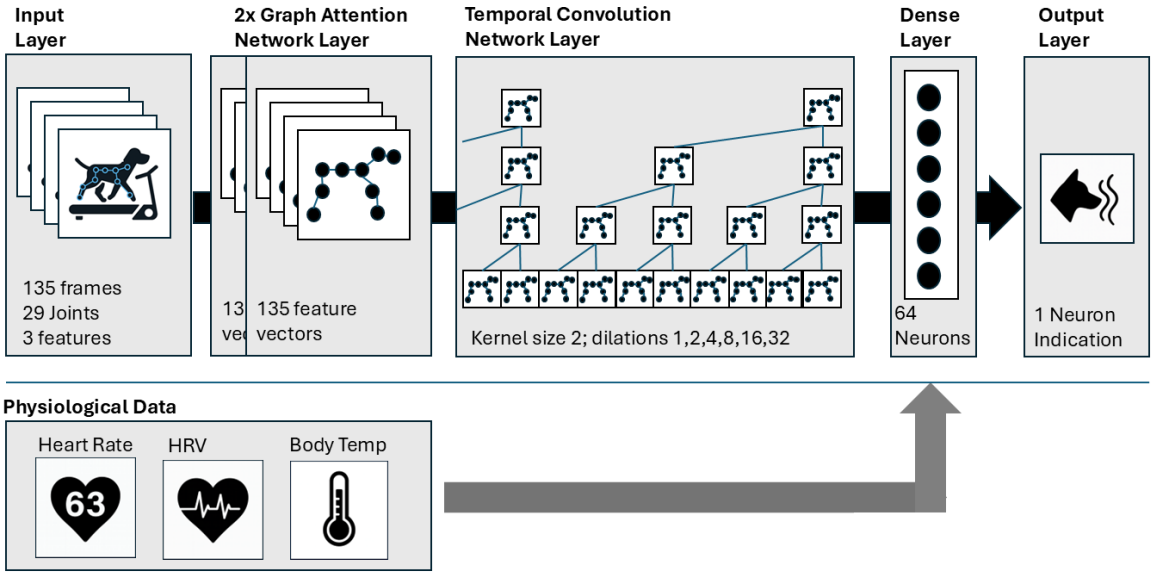

Figure 2. Architecture Spatio-Temporal Network Models. The network processes 135-frame sequences of 29 joints (with x, y, and confidence as features). Pose graphs from each frame are first passed through two stacked Graph Attention Network (GAT) layers, enabling information exchange up to the second-order neighborhood of each joint. These spatial embeddings are then processed by a multi-layer Temporal Convolutional Network (TCN) with increasing dilation factors to capture movement dynamics over time. A final dense layer maps the sequence representation to a binary output predicting whether the dog will indicate a target odor. In an extended version of the model, physiological data (heart rate, heart rate variability, and core temperature) are processed through separate dense layers and fused with the pose-based representation before the final dense layer.

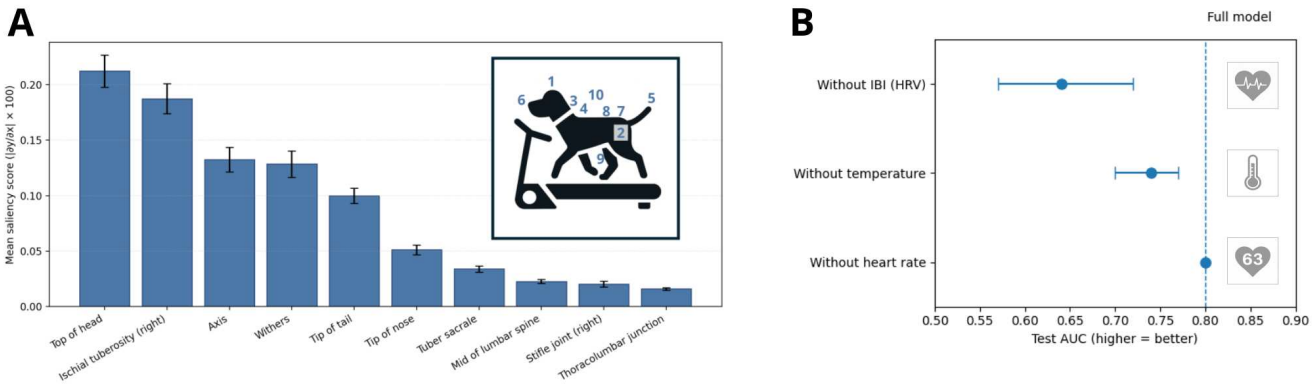

Figure 3. Behavioral and physiological features underlying prediction. (A) Saliency ranking of pose landmarks derived from gradient-based analysis of the video-only model. Mean saliency scores quantify the influence of small perturbations in joint coordinates on predicted miss probability and were averaged across frames and trials; error bars indicate ±SEM. The ten landmarks with the highest saliency scores (out of 29 total landmarks) are shown. Predictive information is concentrated along the axial body line, with cranial (top of head, axis), trunk (withers), and caudal (tail tip) landmarks exerting the strongest influence. The schematic overlay highlights the anatomical locations of the most influential landmarks.
(B) Physiological feature ablation analysis showing model performance after removal of individual modalities. Points indicate test AUC of the multimodal model when one physiological input (inter-beat interval-derived heart rate variability, core temperature, or heart rate) is omitted; horizontal error bars denote 95% confidence intervals. The dashed vertical line marks performance of the full model. Removal of inter-beat interval features results in the largest decrease in predictive performance, indicating that short-term heart rate variability is the dominant physiological contributor to prediction.

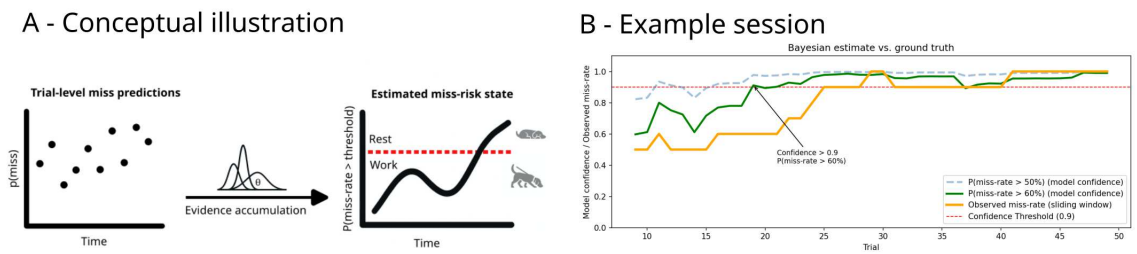

Figure 4. Aggregated model outputs track evolving miss risk within a session. (A) Conceptual illustration. Trial-level predictions of missed alerts derived from pre-odor movement and physiological signals are treated as probabilistic evidence and sequentially integrated to estimate the probability that the dog's underlying miss-rate exceeds a predefined threshold. This transformation shifts interpretation from isolated trial outcomes to a continuous estimate of performance state.
(B) Example session. Sequential aggregation of model outputs across 50 target-present trials from a single dog. The estimated probability that the miss-rate exceeds predefined thresholds (50% and 60%) is shown alongside the observed miss frequency (10-trial sliding window). The aggregated estimate increases as empirical miss frequency rises during the session. The dashed horizontal line indicates a confidence threshold (0.9) for concluding that the dog is operating above the selected miss-rate level. The annotation marks the first point at which the probability that the miss-rate exceeds 60% surpasses this threshold.

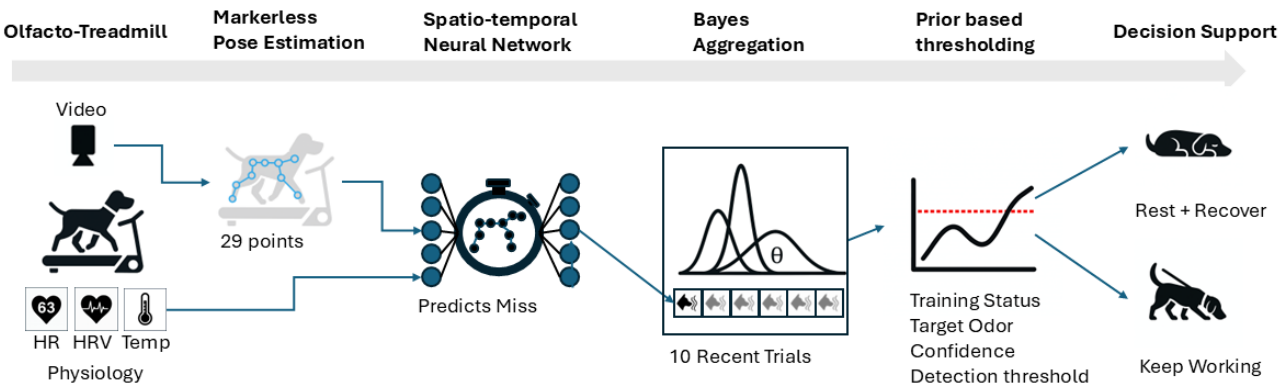

Figure 5. End-to-end system for predicting detection dog performance declines. Video recordings and physiological signals (heart rate, heart rate variability, and core body temperature) are collected during olfacto-treadmill sessions. A markerless pose estimation system extracts movement features, which are combined with physiological input and processed by a spatio-temporal neural network trained to predict missed indications. Model outputs are aggregated over multiple trials using a Bayesian framework to estimate the dog's current likelihood of missing a target odor. A prior-informed thresholding step—adaptable to the training context, target odor, and prediction confidence—then supports decision-making: whether the dog should continue working or be given time to rest and recover.

Tables

Table 1. Demographic information of study dogs.

Dog Name	Sex	Breed	Age (years)	Weight (kg)
Adele	Spayed female	Belgian Malinois	3	22
Boomer	Intact male	English Springer Spaniel	2	23
Charles	Neutered male	Mixed breed	4	20
Iitooma	Neutered male	English Springer Spaniel	3	17

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementaryText.pdf
- S1VideoSegmentedPosePrediction.mp4
- S2VideoSaliency.mp4