EAI Endorsed Transactions
on AI and Robotics
Research Article EAI.EU

Conflict-Weighted Meta-Gradient Correction for Continual Learning in Multi-Sensor UAV Positioning

Nguyen Xuan Truong[1], Tran Thi Van[2], Nguyen Van Thang[3], Quang-Anh N. D.[4] and Khoa Nguyen Dang[4*]

[1]Training Department, Hung Yen University of Technology and Education, Hung Yen, Vietnam
[2]Faculty of Information Technology, University of Labour and Social Affairs, Hanoi, Vietnam
[3]University of Engineering and Technology, Vietnam National University, Hanoi, Vietnam
[4]Faculty of Engineering and Technology, International School, Vietnam National University, Hanoi, Vietnam

Abstract

Multi-sensor UAV positioning systems are increasingly deployed in sequences where new sensor modalities are integrated mid-deployment, requiring a model to expand its input space while retaining knowledge of previously observed environments. Existing continual learning methods assume a fixed-dimensional input and fail in this modality-incremental settings with earlier single-modality task memories produce stale gradients that actively oppose learning of newly introduced sensor features. We propose RAMGCL, a method that addresses this dual challenge of catastrophic forgetting and input-space expansion through three components operating within a fixed-size architecture. An asymmetric modality filter selectively suppresses stale gradients from zero-padded memories while retaining cross-modality signal that remains informative. A conflict-weighted per-task memory gradient continuously corrects the update direction by upweighting tasks whose gradients most conflict with the current update. A bi-level diagonal gradient preconditioner meta-learns a per-dimension scaling vector for the latent gradient slice, calibrating stability and plasticity without solving a quadratic programme. Episodic memories are populated using farthest-point sampling in position space to ensure spatially diverse coverage of each environment. Experiments on the Q-Drone UWB benchmark across five environments and two modality-incremental orderings show that RAMGCL achieves state-of-the-art performance in both average MAE and forgetting, and is the only method to achieve positive backward transfer in the IMU-to-UWB ordering.

Received on 29 May 2026; accepted on 03 November 2026; published on 06 November 2026

Keywords: Continual Learning, Unmanned Aerial Vehicles (UAVs), UAV Positioning, Meta-Gradient

Copyright © 2026 Nguyen Xuan Truong *et al.*, licensed to EAI. This is an open access article distributed under the terms of the CC BY-NC-SA 4.0, which permits copying, redistributing, remixing, transformation, and building upon the material in any medium so long as the original work is properly cited.

doi:10.4108/airo.13236

1. Introduction

Accurate positioning in unconstrained environments is a fundamental requirement for autonomous systems such as Unmanned Aerial Vehicles (UAVs), where robust localisation underpins safe navigation, mission planning, and real-time control [1–3]. UAV platforms equipped with multi-sensor perception have found practical application in domains ranging from field-robot navigation via LiDAR-camera fusion [4] to real-time aerial crop monitoring [5] and sensor-based source localization in robotic olfaction [6], underscoring the breadth of tasks that depend on reliable onboard sensing. Multi-sensor fusion, combining complementary modalities such as Ultra-Wideband (UWB) radio and Inertial Measurement Units (IMU), has emerged as the dominant strategy to improve coverage and resilience, since individual sensors suffer from environment-specific failure modes, as UWB degrades under severe multipath propagation, while IMU drift accumulates unboundedly over time [3, 7]. In practice, however, a deployed positioning system rarely acquires all sensor modalities from the outset. Hardware constraints, cost, or incremental system upgrades mean that new sensors are integrated mid-deployment, requiring the model to

*Corresponding author. Email: khoand@vnuis.edu.vn

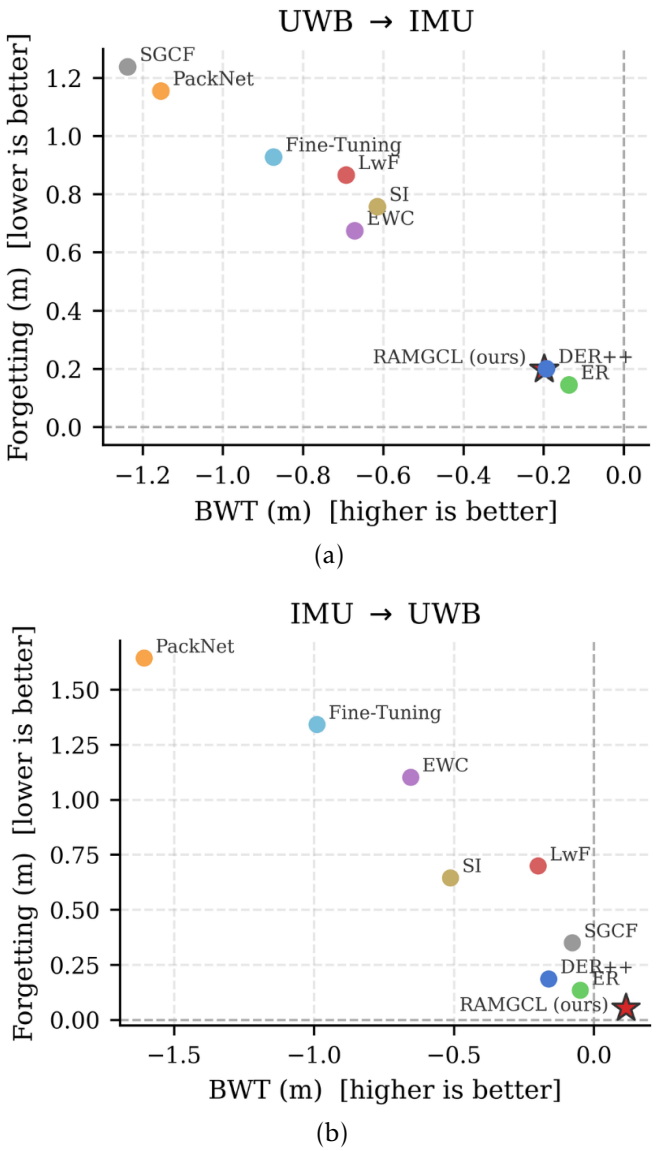


Figure 1. Forgetting vs. backward transfer (BWT) for all compared methods under both MIL orderings: (a) the first MIL ordering; (b) the second MIL ordering. The ideal region (low forgetting, positive BWT) is marked by dashed lines. RAMGCL (star) achieves the best stability-plasticity balance across both directions

extend its input space without forgetting the knowledge accumulated from previously observed environments.

Continual Learning (CL) addresses the problem of learning from a non-stationary stream of tasks without catastrophic forgetting [8, 9]. Existing approaches fall into three broad families with regularisation-based methods, such as Elastic Weight Consolidation (EWC) [10] and Synaptic Intelligence (SI) [11], penalise changes to parameters deemed important for past tasks, but rely on a fixed r egularisation s trength that cannot adapt to the shifting gradient conflict patterns introduced by diverse environments. Knowledge-distillation methods, exemplified by Learning without Forgetting (LwF) [12], preserve past behaviour by matching output distributions, yet require a fixed input dimensionality and become ineffective when new sensor modalities expand the feature space. Gradient projection and memory replay methods, including Gradient Episodic Memory (GEM) [13], A-GEM [14], Dark Experience Replay++ (DER++) [15], and related approaches [16], constrain or correct the current gradient using stored exemplars, offering stronger anti-forgetting guarantees at the cost of either solving a quadratic programme at each step (GEM) or treating all past tasks as equally important (A-GEM). Architecture-based methods such as Progressive Neural Networks [17] and PackNet [18] isolate task-specific capacity but grow the model or require iterative pruning, both of which are impractical for resource-constrained UAV deployment. A critical limitation shared by above methods is the assumption of a fixed-dimensional input space. Under this assumption, zero-padded memories from earlier single-modality tasks are fed unchanged into the model when a new modality is introduced, producing gradients that carry no information about the newly added sensor features. We term these stale gradients, rather than providing protective correction, they actively oppose the model ability to learn the new modality. Beyond stale gradients, the gradient landscapes induced by different environments vary substantially in structure, creating conflict patterns that shift continuously across the task sequence and cannot be handled by a fixed regularisation strength.

These two difficulties arise naturally in the Modality-Incremental Learning (MIL) setting [19], where the input feature space expands as new sensors are introduced mid-deployment. For image-based semantic segmentation, disjoint relevance mapping across modalities has been shown necessary to preserve performance on previously learned sensors when the domain gap between modalities, such as camera, depth, and infrared, is large enough to amplify catastrophic forgetting [20]. A closely related setting arises in road-scene semantic segmentation, where a symmetric LiDAR-RGB architecture trained under a class-incremental schedule requires intra-modality, inter-modality, and cross-modality knowledge distillation applied independently to each modality pathway, since treating the pathways jointly allows one modality to degrade the other after a class-incremental update [21]. The modality-incremental paradigm has also been extended to multimodal large language models through modality-inconsistent CL in which each new task carries a previously unseen modality, CL baselines designed for a fixed modality set transfer poorly to this setting because they conflate forgetting caused by modality

shift with forgetting caused by task-type shift [22]. Collectively, these observations reveal that existing continual learning methods assume a fixed-dimensional regression architecture and therefore cannot naturally accommodate modality-incremental input expansion. To overcome both catastrophic forgetting and the challenges introduced by an expanding feature space, we propose **R**egression-**A**ware **M**eta-**G**radient **C**ontinual **L**earning (**RAMGCL**). RAMGCL combines a bi-level diagonal gradient preconditioner that calibrates per-dimension gradient magnitudes, a conflict-weighted per-task memory gradient that continuously corrects the update direction without solving a quadratic programme, and an asymmetric modality-aware memory filter that suppresses stale gradients while retaining cross-modality signal. Together, these components achieve low catastrophic forgetting and strong forward transfer across both MIL directions within a fixed-size architecture.

The main contributions of this paper are summarised as follows:

- We propose RAMGCL, a lightweight method that combines a bi-level gradient preconditioner, conflict-weighted memory correction, and an asymmetric modality filter, all within a fixed-size architecture.
- We introduce spatially diverse episodic memory via farthest-point sampling in position space, ensuring anchor sets cover the full spatial extent of each environment without manual curation.
- We conduct comprehensive experiments on a real-world multi-environment UAV positioning dataset across five environments, demonstrating that RAMGCL achieves state-of-the-art performance under both MIL orderings with low catastrophic forgetting.

The remainder of this paper is organised as follows. Section 2 formalises the MIL problem setting and presents the RAMGCL method, detailing the bi-level diagonal gradient preconditioner, the conflict-weighted per-task memory gradient, the asymmetric modality-aware memory filter, and the spatially diverse episodic memory construction via farthest-point sampling. Section 3 describes the Q-Drone UWB benchmark and implementation details, and reports comparative results against eight continual learning baselines under both MIL orderings, along with component and hyperparameter ablation studies. Section 4 concludes the paper and discusses limitations and directions for future work.

2. Methodology

2.1. Problem Setup

Multi-sensor positioning in unconstrained environments is inherently a sequential learning problem. A deployed system encounters new environments one after another and, in many practical settings, the set of available sensors changes across this sequence: an initial deployment phase may rely on a single modality, after which additional sensors are integrated mid-deployment. We formalise this scenario as *modality-incremental continual learning* (MIL), a specialisation of task-incremental CL [19] in which the input feature space expands when a new modality is introduced, posing challenges that fixed-input CL methods cannot directly address.

Formal definition. Let $\mathcal{T} = \{(\mathcal{D}_t, m_t)\}_{t=1}^{T}$ be an ordered task sequence, where $\mathcal{D}_t = \{(\mathbf{x}_i^t, \mathbf{y}_i^t)\}_{i=1}^{N_t}$ is the dataset for task t and $m_t \in \mathcal{M}_s$ denotes the active sensor modality from a finite modality set $\mathcal{M}_s = \{\mu_1, \mu_2, \mu_{12}\}$, where μ_1 and μ_2 denote the two single-sensor modalities and μ_{12} denotes their fusion. Each input $\mathbf{x}_i^t \in \mathbb{R}^{d_{m_t}}$ is a windowed feature vector extracted from raw sensor streams, and $\mathbf{y}_i^t \in \mathbb{R}^3$ is the 3D position in metres. Feature dimensionalities are d_{μ_1}, d_{μ_2}, and $d_{\mu_{12}} = d_{\mu_1} + d_{\mu_2}$ for the single and fused modalities respectively, in our UAV positioning instantiation, $d_{\mu_1} \in \{20, 200\}$ depending on direction and $d_{\mu_{12}} = 220$. All inputs are zero-padded to the unified maximum width $d = d_{\mu_{12}}$ so that a single regression model $f_{\boldsymbol{\theta}} : \mathbb{R}^d \to \mathbb{R}^3$ operates throughout the entire sequence without architectural modification. At step t only $\mathcal{D}_t$ is available, raw data from previous tasks cannot be replayed beyond a compact episodic memory buffer $\mathcal{M} \subset \mathbb{R}^d \times \mathbb{R}^3$. The global objective is to minimise mean absolute error (MAE) across all tasks simultaneously:

$$\min_{\boldsymbol{\theta}} \frac{1}{T} \sum_{t=1}^{T} \mathbb{E}_{(\mathbf{x},\mathbf{y})\sim\mathcal{D}_t} \Big[\| f_{\boldsymbol{\theta}}(\mathbf{x}) - \mathbf{y} \|_1 \Big], \tag{1}$$

subject to the constraint that only $\mathcal{D}_t$ and $\mathcal{M}$ may be accessed when learning task t.

Challenges specific to MIL. Standard CL methods address forgetting through parameter regularisation [10, 11], knowledge distillation [12], or gradient projection [13]. All of these assume a fixed-dimensional input space and treat every gradient dimension as equally significant. The MIL setting introduces two additional difficulties. First, when modality μ_{12} is introduced after μ_1, the stored μ_1-memories have zeros in the feature dimensions belonging to μ_2. Backpropagating through these zero-padded inputs produces gradients that carry no information about the newly introduced sensor features, and can actively oppose the model's

ability to learn them. We call these *stale gradients* and resolve them in Section 2.2. Second, the gradient landscapes induced by different environments vary substantially in structure because each environment has distinct spatial coverage, trajectory density, and sensor measurement characteristics, creating gradient conflict patterns that shift continuously across the task sequence and cannot be handled by a fixed regularisation strength. These two difficulties jointly motivate RAMGCL.

2.2. Proposed RAMGCL

Backbone and Notation. Sensor-based positioning tasks of moderate complexity can be solved effectively with simple feedforward architectures, prior work on UWB and IMU localisation consistently shows that compact MLPs match or exceed more complex architectures when the feature extraction is carefully designed [3]. Accordingly, we design a lightweight shared backbone $f_\theta : \mathbb{R}^d \to \mathbb{R}^3$ as a three-layer MLP with width sequence $d \mapsto 256 \mapsto d_\ell \mapsto 3$, where $d_\ell = 128$ is the latent dimension, with ReLU activations and Layer Norm applied after the first projection. This architecture sufficient enough to represent the position-dependent signal patterns of all five environments while remaining small enough for the meta-learning components of RAMGCL to converge reliably.

We vectorise all learnable parameters as $\theta \in \mathbb{R}^p$ ($p = |\theta|$) and write ∇_θ for the corresponding gradient operator. The task loss at step t uses mean squared error: $\mathcal{L}_t(\theta) = \mathcal{L}_{\text{MSE}}(f_\theta, \mathcal{D}_t)$. The episodic memory buffer is $\mathcal{M} = \{(\mathbf{X}_j, \mathbf{Y}_j, m_j)\}_{j=1}^{t-1}$, where $\mathbf{X}_j \in \mathbb{R}^{B_{\text{anc}} \times d}$ and $\mathbf{Y}_j \in \mathbb{R}^{B_{\text{anc}} \times 3}$ hold the B_{anc} anchor samples for past task j, and $m_j \in \mathcal{M}_s$ records its modality label. The core meta-parameter is a *diagonal gradient preconditioner* $\mathbf{d} \in \mathbb{R}^{d_\ell}$, constrained to the latent-dimension slice of the parameter vector. Given any full gradient $\mathbf{g} \in \mathbb{R}^p$, the preconditioned gradient is defined as:

$$\tilde{\mathbf{g}} \;=\; \Big[\, \mathbf{d} \odot \mathbf{g}_{1:d_\ell} \;\Big|\; \mathbf{g}_{d_\ell+1:p} \,\Big], \tag{2}$$

where $\odot$ denotes element-wise multiplication and $[\,\cdot\,|\,\cdot\,]$ denotes concatenation. Restricting $\mathbf{d}$ to the latent slice $\mathbf{g}_{1:d_\ell}$ keeps the number of meta-parameters at $d_\ell = 128 \ll p$, concentrating calibration on the representation layer that is most shared across modalities and most susceptible to cross-task interference.

Bi-level Gradient Preconditioner. The preconditioner $\mathbf{d}$ is trained by a bi-level procedure that measures how well a preconditioned gradient step preserves past task performance, and backpropagates this signal into $\mathbf{d}$. Every τ_{meta} training steps and whenever $\mathcal{M} \neq \emptyset$, the following two-phase update is performed.

Inner phase. A batch $(\mathbf{X}_\mathcal{M}, \mathbf{Y}_\mathcal{M})$ is sampled uniformly from all stored memory entries. The current task gradient on this batch is computed and preconditioned to obtain $\tilde{\mathbf{g}}_\mathcal{M}$ via Eq. (2). A *virtual* parameter update is then applied:

$$\theta' \;=\; \theta - \eta\, \tilde{\mathbf{g}}_\mathcal{M}, \tag{3}$$

where η is the main learning rate. The key word is *virtual*, θ is stored before the step and will be restored afterwards, so this phase produces a candidate parameter θ' solely for evaluation purposes.

Outer phase. The quality of the candidate θ' is assessed by an *MAE meta-loss* evaluated on the same memory batch:

$$\mathcal{L}_{\text{meta}}(\mathbf{d}) \;=\; \frac{1}{|\mathcal{M}_{\text{batch}}|} \sum_i \big\| f_{\theta'}(\mathbf{x}_i^\mathcal{M}) - \mathbf{y}_i^\mathcal{M} \big\|_1. \tag{4}$$

The gradient $\nabla_\mathbf{d} \mathcal{L}_{\text{meta}}$ flows through the virtual step in Eq. (3) because θ' depends on $\mathbf{d}$ via $\tilde{\mathbf{g}}_\mathcal{M}$. The preconditioner is then updated and projected back onto its feasible set:

$$\mathbf{d} \;\leftarrow\; \Pi_{[d_{\min}, d_{\max}]}\Big(\mathbf{d} - \eta_{\text{meta}}\, \nabla_\mathbf{d} \mathcal{L}_{\text{meta}} \Big), \tag{5}$$

where $\Pi_{[d_{\min}, d_{\max}]}$ clips element-wise to $[d_{\min}, d_{\max}] = [0.1, 10.0]$. After Eq. (5), θ is restored to its pre-inner-step value, so only $\mathbf{d}$ is modified by the outer loop. This restoration guarantees that the meta-update is a pure scale-calibration step with no side-effect on the main learning trajectory, making's behaviour stable and predictable across different task orderings.

Conflict-Weighted Per-Task Memory Gradient. The meta-update in Section 2.2 refines $\mathbf{d}$ every τ_{meta} steps, but between these updates uncorrected gradients from the current task can drift the model away from past performance. To address this, we augment every training step with a lightweight *per-task memory gradient correction* that provides continuous directional protection without solving a quadratic programme.

For each past task $j \in \{1, \ldots, t-1\}$, a dedicated reference gradient is computed separately from the B_{mem} anchor samples of that task, rather than pooling all memories into a single batch. Pooling would average out the per-task conflict signal and cause older tasks to receive systematically less protection as the buffer grows. The per-task gradient is:

$$\mathbf{g}_j^\mathcal{M} \;=\; \nabla_\theta\, \mathcal{L}_{\text{MSE}}\Big(f_\theta(\mathbf{X}_j^\mathcal{M}), \mathbf{Y}_j^\mathcal{M} \Big). \tag{6}$$

Each per-task gradient $\mathbf{g}_j^\mathcal{M}$ is assigned a conflict weight w_j that measures how strongly task j conflicts with the current task gradient $\mathbf{g}_t$:

$$w_j \;=\; \max\Big(0,\; -\cos(\mathbf{g}_t, \mathbf{g}_j^\mathcal{M})\Big) + \beta, \tag{7}$$

where $\cos(\mathbf{u}, \mathbf{v}) = \mathbf{u}^\top \mathbf{v} / (\|\mathbf{u}\|_2 \|\mathbf{v}\|_2 + \varepsilon)$ is the cosine similarity and $\beta = \max(1/|\mathcal{M}|,\, 0.5)$ is a floor weight. The

first term in Eq. (7) assigns extra weight to tasks whose gradient actively conflicts with the current update direction, following the intuition that these tasks are at greatest risk of being forgotten. The conflict-weighted aggregate memory gradient is then:

$$\mathbf{g}^{\mathcal{M}} = \frac{\sum_j w_j \, \mathbf{g}_j^{\mathcal{M}}}{\sum_j w_j}. \tag{8}$$

At each training step, the current task gradient $\mathbf{g}_t$ and the memory correction $\mathbf{g}^{\mathcal{M}}$ are combined with mixing coefficient $\alpha \in (0, 1)$:

$$\hat{\mathbf{g}}_t = \mathbf{g}_t + \alpha \, \mathbf{g}^{\mathcal{M}}, \tag{9}$$

and the combined gradient is preconditioned via Eq. (2) before the parameter update:

$$\theta \leftarrow \theta - \eta \, \widetilde{\hat{\mathbf{g}}}_t. \tag{10}$$

Asymmetric Modality-Aware Memory Filtering. The memory gradient correction in Section 2.2 would be compromised by the stale gradients identified in Section 2.1 if all stored memories were used indiscriminately. Recall that when the sequence transitions from single-modality tasks ($m_j = \mu_1$) to fused tasks ($m_t = \mu_{12}$), the memories stored for μ_1-tasks have zeros in the feature dimensions corresponding to μ_2 (indices $d_{\mu_1} + 1$ through $d_{\mu_{12}}$). Computing $\mathbf{g}_j^{\mathcal{M}}$ from these inputs yields gradients that carry no signal about μ_2 features, actively resisting the model's ability to learn the newly introduced sensor.

To eliminate this interference without discarding all cross-modality protection, we apply asymmetric modality filtering. Let $m_{\text{first}} = m_1$ be the modality of the first stored task, which encodes the direction of the MIL sequence. The *active* memory subset used for per-task gradient computation is defined as:

$$\mathcal{M}_{\text{active}} = \begin{cases} \{j : m_j = \mu_{12}\} & \text{if } m_t = \mu_{12} \text{ and } m_{\text{first}} = \mu_1, \\ \mathcal{M} & \text{otherwise.} \end{cases} \tag{11}$$

The asymmetry in Eq. (11) is deliberate and reflects a structural property of the zero-padding scheme. When the sequence proceeds as $\mu_1 \to \mu_{12}$ (sensor 1 first, then fusion), the early μ_1-memories have zeros in the μ_2 feature block; their gradients are stale with respect to the fused model and are therefore excluded. When the sequence proceeds as $\mu_2 \to \mu_{12}$ (sensor 2 first, then fusion), the early μ_2-memories have zeros in the μ_1 feature block, but the μ_1 component of the fused input (indices 1 through d_{μ_1}) is populated with real sensor readings. The gradient from a μ_2-memory therefore still provides valid signal for the μ_1 feature dimensions and is retained for protection. This asymmetric treatment encodes domain knowledge about sensor feature layout as a structural inductive bias, a design principle not present in any general-purpose CL method.

Spatially Diverse Episodic Memory. All gradient-based memory mechanisms in RAMGCL depend on the stored anchor samples $\mathbf{X}_j^{\mathcal{M}}$ being representative of the full spatial extent of each past environment. Trajectory data from positioning tasks is highly non-uniform: frequently traversed corridors or repeated waypoints accumulate many samples, whereas the extremes of the environment may be visited only once. Selecting anchor samples uniformly at random risks over-representing the dense regions and leaving the sparse ones unprotected.

We address this by selecting B_{anc} anchor samples per completed task using Farthest-Point Sampling (FPS) [23] in the ground-truth position space $\mathcal{Y} \subseteq \mathbb{R}^3$. The selected set $\mathcal{S}^* \subseteq \mathcal{D}_t$ maximises the minimum pairwise Euclidean distance among selected positions:

$$\mathcal{S}^* = \underset{\substack{\mathcal{S} \subseteq \mathcal{D}_t \\ |\mathcal{S}| = B_{\text{anc}}}}{\arg\max} \; \min_{\substack{i,k \in \mathcal{S} \\ i \neq k}} \|\mathbf{y}_i - \mathbf{y}_k\|_2. \tag{12}$$

The FPS algorithm solves Eq. (12) greedily in $O(B_{\text{anc}} \cdot |\mathcal{D}_t|)$ time, which is negligible compared with training. Alongside each anchor set $(\mathbf{X}_j^{\mathcal{M}}, \mathbf{Y}_j^{\mathcal{M}})$, the modality label m_j is stored to enable the filter in Eq. (11). The buffer grows by one entry per task at a fixed cost of B_{anc} samples, keeping total memory overhead constant regardless of sequence length. Furthermore, Algorithm 1 outlines the pseudo calculation of the proposed method.

3. Experiments

3.1. Dataset

We evaluate RAMGCL on the Q-Drone UWB benchmark [3], a publicly available dataset collected with a custom UAV platform equipped with a PulsON 440 UWB tag and an onboard IMU. The platform was flown across five environments that differ substantially in spatial coverage, structural complexity, and multipath conditions: an indoor warehouse, an open sports field, a site close to a glass-facade building, the underside of a concrete-metal bridge, and a semi-open metal tunnel. Table 1 summarises the key characteristics of each site.

Sensor data were collected over 23 flight sessions totalling approximately 1 hour 50 minutes and 4.3 km of travelled distance. Each session provides time-synchronised UWB ranging measurements from four ground anchors, raw IMU readings (linear acceleration, angular velocity, and barometric height), and 3D positions independently measured by robotic total stations tracking a mini prism mounted on the UAV.

Algorithm 1 Training procedure of RAMGCL.

Require: Task sequence $\{(\mathcal{D}_t, m_t)\}_{t=1}^{T}$, model f_θ
Require: η, η_{meta}, τ_{meta}, α, β, $d_{\min}$, $d_{\max}$, B_{anc}, B_{mem}
$\mathbf{d} \leftarrow \mathbf{1}_{d_\ell}$; $\mathcal{M} \leftarrow \emptyset$; $k \leftarrow 0$; $m_{\text{first}} \leftarrow m_1$
for $t = 1, \ldots, T$ **do**
  for each mini-batch $(\mathbf{x}, \mathbf{y}) \sim \mathcal{D}_t$ **do**
    $k \leftarrow k + 1$
    if $k \bmod \tau_{\text{meta}} = 0$ and $\mathcal{M} \neq \emptyset$ **then**
      $\theta_{\text{save}} \leftarrow \theta$
      $\theta' \leftarrow \theta - \eta\, \tilde{\mathbf{g}}_{\mathcal{M}}$ *(Eq. 3)*
      $\mathbf{d} \leftarrow \Pi_{[d_{\min}, d_{\max}]}(\mathbf{d} - \eta_{\text{meta}} \nabla_{\mathbf{d}} \mathcal{L}_{\text{meta}})$ *(Eqs. 4, 5)*
      $\theta \leftarrow \theta_{\text{save}}$
    end if
    $\mathbf{g}_t \leftarrow \nabla_\theta \mathcal{L}_t(\theta)$
    if $\mathcal{M} \neq \emptyset$ **then**
      Determine $\mathcal{M}_{\text{active}}$ *(Eq. 11)*
      Compute $\{\mathbf{g}_j^{\mathcal{M}}\}_{j \in \mathcal{M}_{\text{active}}}$ *(Eq. 6)*
      Compute $\{w_j\}$ and $\mathbf{g}^{\mathcal{M}} \leftarrow \sum_j w_j \mathbf{g}_j^{\mathcal{M}} / \sum_j w_j$ *(Eqs. 7, 8)*
      $\hat{\mathbf{g}}_t \leftarrow \mathbf{g}_t + \alpha\, \mathbf{g}^{\mathcal{M}}$ *(Eq. 9)*
    else
      $\hat{\mathbf{g}}_t \leftarrow \mathbf{g}_t$
    end if
    $\theta \leftarrow \theta - \eta\, \widetilde{\hat{\mathbf{g}}}_t$ *(Eq. 10)*
  end for
  $\mathcal{M} \leftarrow \mathcal{M} \cup \{(\mathbf{X}_t^{\text{fps}}, \mathbf{Y}_t^{\text{fps}}, m_t)\}$ *(Eq. 12)*
end for

Table 1. Q-Drone benchmark dataset summary [3]

Site	Setting	Area (m^2)
Indoor	Indoor	91.96
Field	Outdoor, open	158.74
Building	Outdoor, near building	50.10
Bridge	Semi-open, under bridge	246.80
Tunnel	Semi-open, tunnel	241.80
Total		789.40

3.2. Implementation Details

All experiments are implemented on a single NVIDIA RTX4070 SUPER GPU. Both θ and $\mathbf{d}$ are optimised with Adam at learning rates $\eta = 10^{-3}$ and $\eta_{\text{meta}} = 10^{-4}$ respectively. The clip bounds are $d_{\min} = 0.1$ and $d_{\max} = 10.0$; the mixing coefficient is $\alpha = 0.5$, the floor weight is $\beta = 0.5$. Each task trains for at most 30 epochs with early stopping at patience 7 on the held-out validation MAE. The meta-update period is $\tau_{\text{meta}} = 20$ steps, using a memory batch of 64 samples drawn uniformly from $\mathcal{M}$. Per-task memory gradients use $B_{\text{mem}} = 16$ anchor samples each, the FPS buffer stores $B_{\text{anc}} = 16$ anchor samples per task. Sensor features are extracted with sliding windows of width $W_{\mu_1} = 20$ (sensor 1) and $W_{\mu_2} = 10$ (sensor 2) at stride $s = 5$, capped at $N_{\max} = 2{,}000$ samples per environment.

3.3. Evaluation Results

Table 2 reports the per-environment MAE of models trained on a single modality in isolation, establishing an upper bound on difficulty for each environment. UWB-only achieves a mean MAE of 2.583 m, while IMU-only reaches 3.251 m, reflecting t he c omplementary failure modes of the two sensors, UWB degrades in structurally complex sites (Field: 4.066 m; Tunnel: 4.339 m) where multipath and signal blockage are severe, whereas IMU suffers f rom u nbounded d rift i n t he Tunnel environment (7.722 m). These results motivate sensor fusion and confirm that no single modality is sufficient across all five environments.

Table 2. Single-modality baseline MAE per environment. Lower MAE indicates better performance

Setting	Indoor	Field	Building	Bridge	Tunnel	Mean
UWB only	0.889	4.066	2.029	1.593	4.339	2.583
IMU only	1.188	4.264	1.135	1.945	7.722	3.251

Table 3 and Figure 2 compare RAMGCL against eight CL methods under both MIL orderings. RAMGCL achieves the lowest Avg MAE in both directions (2.328 m and 2.398 m) and the lowest forgetting (0.199 and 0.054). Regularisation-based methods (EWC, SI, LwF) suffer high forgetting across both orderings, as their fixed importance weights cannot adapt to the shifting gradient structure of heterogeneous environments, and their fixed-dimensionality assumptions are directly violated when a new modality expands the input space. PackNet incurs the largest forgetting of any method (up to 1.643), indicating that iterative pruning produces subnetworks too fragile for modality shift. Among replay-based methods, DERpp and ER outperform the regularisation family, confirming that episodic memory provides a stronger anti-forgetting signal in this setting, RAMGCL improves further by replacing uniform memory gradients with conflict-weighted per-task corrections and by suppressing stale gradients from zero-padded memories. The gain is most pronounced in the IMU+UWB direction, where RAMGCL is the only method to achieve positive BWT (0.114), meaning that later tasks marginally improve rather than degrade performance on earlier ones. SGCF achieves the highest FWT in UWB+IMU (1.340) but at the cost of the highest Avg MAE (3.187 m) and forgetting (1.237) in that direction, suggesting its gradient orthogonality constraint over-constrains the update. Per-environment results in Table 4 show that the Field site remains the hardest in both orderings due to its large spatial extent and sparse anchor coverage, while the Tunnel directional gap (2.733 vs. 1.967 m) reflects the asymmetric filter, under UWB+IMU, stored UWB memories continue providing valid gradient correction

Table 3. Comparison of CL methods under different settings. Lower Avg MAE and Forgetting are better, while higher BWT and FWT are preferred

Method	UWB + IMU				IMU + UWB			
	Avg MAE ↓	BWT ↑	FWT ↑	Forget ↓	Avg MAE ↓	BWT ↑	FWT ↑	Forget ↓
DERpp	2.341	-0.193	0.861	0.200	2.662	-0.161	0.805	0.185
ER	2.529	-0.136	0.793	0.144	2.575	-0.048	1.131	0.133
FineTuning	2.839	-0.873	0.948	0.927	3.062	-0.989	0.764	1.341
LwF	2.686	-0.692	0.634	0.865	2.614	-0.199	0.731	0.699
PackNet	2.934	-1.154	1.071	1.154	3.689	-1.607	0.844	1.643
EWC	3.166	-0.671	0.451	0.674	2.896	-0.654	0.840	1.101
SI	2.781	-0.614	0.628	0.757	2.764	-0.512	1.142	0.644
SGCF	3.187	-1.237	**1.340**	1.237	2.505	-0.077	0.522	0.349
RAMGCL	**2.328**	**-0.199**	0.767	**0.199**	**2.398**	**0.114**	0.829	**0.054**

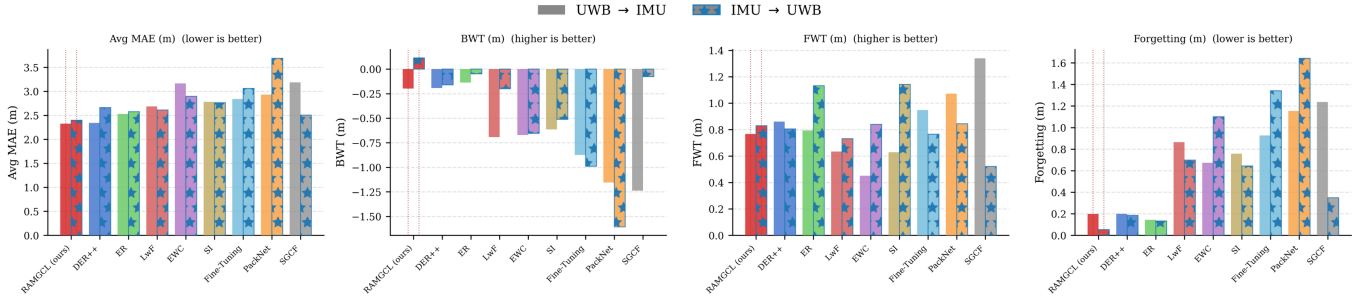


Figure 2. Comparison of all CL methods on four metrics under both MIL orderings

after modality expansion, whereas the reverse ordering must discard early memories to avoid stale gradients.

Table 4. Per-environment MAE of RAMGCL. Lower MAE indicates better performance

Setting	Indoor	Field	Building	Bridge	Tunnel	Mean
IMU + UWB	1.106	3.985	1.846	2.321	2.733	2.398
UWB + IMU	1.067	4.272	1.893	2.441	1.967	2.328

3.4. Ablation Study

Architecture Ablations. Table 5 reports leave-one-out component ablations. Removing FPS memory causes the largest degradation overall (UWB+IMU: MAE 2.833 m, forgetting 0.740), confirming that spatially diverse anchor selection is critical for representative gradient correction, random sampling leaves sparse environments under-protected and dense ones over-represented. Removing the preconditioner degrades both orderings, with the strongest effect in IMU+UWB (MAE 2.708 m, forgetting 0.470) where the larger feature expansion produces stronger gradient interference, and a notable forgetting increase under UWB+IMU as well (0.196 vs. 0.199), confirming that latent-slice calibration is necessary in both directions. Removing conflict weighting raises forgetting substantially under UWB+IMU (0.380 vs. 0.199), as uniform memory gradients give equal weight to at-risk and already-stable tasks. Removing the modality filter degrades MAE under IMU+UWB (2.483 m), consistent with the stale gradient analysis: without filtering, zero-padded μ_1 memories pollute the gradient correction for the fused modality. Notably, removing the memory gradient has minimal impact under IMU+UWB (MAE 2.315 m, forgetting 0.062), suggesting that the preconditioner and filter together are sufficient for stability in that direction and that the memory gradient contributes primarily under UWB+IMU where gradient conflicts are more pronounced.

Hyperparameter Ablations. Table 6 shows that RAMGCL is broadly robust across hyperparameter choices, with the default configuration consistently achieving the lowest forgetting under IMU+UWB (0.054) regardless of the variant tested. Removing the conflict floor ($\beta = 0.0$) causes the largest MAE degradation in both directions (2.651 and 2.518 m), as tasks with near-zero cosine conflict receive no corrective gradient; higher values of β gradually recover MAE but RAMGCL ($\beta = 0.5$) retains the best forgetting balance. The memory mixing coefficient

Table 5. Component ablation of RAMGCL under both MIL orderings. Lower MAE and Forgetting are better

	UWB + IMU		IMU + UWB	
Configuration	**MAE ↓**	**Fgt ↓**	**MAE ↓**	**Fgt ↓**
w/o Preconditioner	2.341	0.196	2.708	0.470
w/o Memory gradient	2.340	0.224	2.315	0.062
w/o Conflict weighting	2.394	0.380	2.319	0.073
w/o Modality filter	2.312	0.185	2.483	0.086
w/o FPS memory	2.833	0.740	2.349	0.162
RAMGCL	**2.328**	**0.199**	**2.398**	**0.054**

α shows a monotonic trend: lower α reduces the memory correction signal and raises both MAE and forgetting, while $\alpha = 1.0$ improves MAE slightly but RAMGCL still achieves lower forgetting in the IMU+UWB direction (0.054 vs. 0.121), indicating that the default mixing ratio better preserves past task performance. Both alternative meta-update periods degrade performance relative to $\tau_{\text{meta}} = 20$, with the slower period ($\tau_{\text{meta}} = 50$) consistently worse across all metrics, confirming that the default frequency provides sufficient preconditioner adaptation without overfitting to the current task. Reducing B_{mem} to 8 raises forgetting notably (UWB+IMU: 0.298), as noisy per-task gradient estimates weaken conflict weighting; increasing to 32 improves MAE but RAMGCL with 16 samples achieves the best forgetting overall.

Table 6. Hyperparameter sensitivity of RAMGCL. Lower MAE and Forgetting are better.

		UWB + IMU		IMU + UWB	
Param	**Value**	**MAE ↓**	**Fgt ↓**	**MAE ↓**	**Fgt ↓**
β	0.0	2.651	0.243	2.518	0.198
	0.25	2.392	0.221	2.447	0.164
	1.0	2.321	0.208	2.421	0.133
α	0.0	2.488	0.267	2.541	0.201
	0.3	2.401	0.214	2.463	0.146
	1.0	2.366	0.203	2.417	0.121
τ_{meta}	10	2.394	0.228	2.481	0.172
	50	2.446	0.251	2.529	0.187
B_{mem}	8	2.517	0.298	2.483	0.191
	32	2.396	0.217	2.438	0.148
RAMGCL		**2.328**	**0.199**	**2.398**	**0.054**

Computational overhead. Table 7 reports per-step wall-clock times for each RAMGCL component, the conflict-weighted memory gradient (B) adds approximately 1.5 ms per past task per step, scaling linearly with T. The bi-level meta-update (C) costs 7.9 ms raw but fires only every τ_{meta}=20 steps, giving an amortised cost of 0.39 ms per step that remains flat across T. At the worst case (T=4), both components together add 2.0 s of overhead per task across 310 gradient steps. At deployment, RAMGCL performs a single forward pass with no memory access, yielding an inference latency of 0.55 ms, satisfying positioning budgets at up to 100 Hz.

Table 7. Per-step wall-clock overhead

Component	ms/step	Per task
A. Baseline (standard step)	2.62	-
B. + Conflict-weighted memory gradient	8.75	1.9 s
C. + Meta-update (amortised ÷20)	0.39	0.1 s
B+C combined	6.52	2.0 s
D. Inference (single query, no gradient)	0.55 ms	

4. Conclusion

We presented RAMGCL, a CL method for multi-sensor UAV positioning that jointly addresses catastrophic forgetting and modality-incremental input expansion within a fixed-size architecture. The core insight is that standard replay-based methods are undermined in this setting by stale gradients produced when zero-padded memories from earlier single-modality tasks are used after a new modality is introduced. RAMGCL resolves this through three components, an asymmetric modality filter that suppresses stale gradients while retaining cross-modality signal, a conflict-weighted memory gradient that prioritises correction of the most at-risk past tasks, and a bi-level diagonal preconditioner that meta-learns per-dimension gradient scaling concentrated on the shared latent representation. Experiments on the Q-Drone benchmark across five environments demonstrate that RAMGCL achieves the best average MAE and forgetting under both MIL orderings.

Limitation. Several limitations of the current work point to future directions. First, all experiments are conducted on a single benchmark dataset, evaluating RAMGCL on additional multi-sensor positioning datasets would strengthen the generalisability of the findings. Second, the current study does not report inference latency or computational overhead under real deployment conditions, as experiments were conducted offline on a single GPU. Profiling RAMGCL on resource-constrained onboard hardware and assessing whether the bi-level meta-update overhead is acceptable for real-time operation are important steps before practical deployment. The task sequence length and modality set are assumed fixed, extending to open-ended sequences where future modalities are unknown, or to settings where sensors are removed as well as added, remains an open direction for future work.

The MLP backbone processes pre-aggregated windowed features and does not model longer-range temporal dependencies, since the target output is an instantaneous 3D position rather than a sequence, this is sufficient for the current benchmark, however replacing the backbone with an LSTM or Transformer encoder while preserving the RAMGCL continual learning mechanism is a natural direction for future work on tasks where raw sensor sequences span longer time horizons.

Declaration of competing interest

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.

References

[1] Moshayedi AJ, Roy AS, Karan U, Zhang M, Bassir D. An integrated beetle antennae search-enabled navigation framework for omnidirectional AGV mobile robots in unknown environments. EAI Endorsed Trans AI Robot. 2026;5.

[2] Li W, Yang F, Li Z, Mou J, Ha Y, Liu Y, et al. A highly efficient ACO-SA algorithm for robot path planning. EAI Endorsed Trans AI Robot. 2025;5.

[3] Arjmandi Z, Kang J, Park K, Sohn G. Benchmark dataset of ultra-wideband radio based UAV positioning. In: Proceedings of the 2020 IEEE 23rd International Conference on Intelligent Transportation Systems (ITSC); 2020 Sep; pp. 1–8.

[4] Moshayedi AJ, Khan AS, Yang Y, Hu J, Kolahdooz A. Robots in agriculture: revolutionizing farming practices. EAI Endorsed Trans AI Robot. 2024;3.

[5] Khan AT, Jensen SM, Khan N. Advancing food security through precision agriculture: YOLOv8's role in efficient pest detection and management. EAI Endorsed Trans AI Robot. 2025;1(5).

[6] Moshayedi AJ, Khan AS, Shuxin Y, Kuan G, Jiandong H, Soleimani M, et al. E-nose design and structures from statistical analysis to application in robotics: a comprehensive review. EAI Endorsed Trans AI Robot. 2023;2(1):e1.

[7] Tian Q, Wang KI-K, Salcic Z. An INS and UWB fusion-based gyroscope drift correction approach for indoor pedestrian tracking. Sensors. 2020;20(16):4476.

[8] McCloskey M, Cohen NJ. Catastrophic interference in connectionist networks: the sequential learning problem. In: Psychology of learning and motivation. Vol. 24. Academic Press; 1989. p. 109–165.

[9] Parisi GI, Kemker R, Part JL, Kanan C, Wermter S. Continual lifelong learning with neural networks: a review. Neural Netw. 2019;113:54–71.

[10] Kirkpatrick J, Pascanu R, Rabinowitz N, Veness J, Desjardins G, Rusu AA, et al. Overcoming catastrophic forgetting in neural networks. Proc Natl Acad Sci U S A. 2017;114(13):3521–3526.

[11] Zenke F, Poole B, Ganguli S. Continual learning through synaptic intelligence. In: Proceedings of the International Conference on Machine Learning; 2017. p. 3987–3995.

[12] Li Z, Hoiem D. Learning without forgetting. IEEE Trans Pattern Anal Mach Intell. 2018;40(12):2935–2947.

[13] Lopez-Paz D, Ranzato MA. Gradient episodic memory for continual learning. In: Advances in Neural Information Processing Systems. 2017;30.

[14] Chaudhry A, Ranzato MA, Rohrbach M, Elhoseiny M. Efficient lifelong learning with A-GEM. arXiv preprint arXiv:1812.00420. 2018.

[15] Buzzega P, Boschini M, Porrello A, Abati D, Calderara S. Dark experience for general continual learning: a strong, simple baseline. In: Advances in Neural Information Processing Systems. 2020;33:15920–15930.

[16] Chaudhry A, Rohrbach M, Elhoseiny M, Ajanthan T, Dokania PK, Torr PH, et al. On tiny episodic memories in continual learning. arXiv preprint arXiv:1902.10486. 2019.

[17] Rusu AA, Rabinowitz NC, Desjardins G, Soyer H, Kirkpatrick J, Kavukcuoglu K, et al. Progressive neural networks. arXiv preprint arXiv:1606.04671. 2016.

[18] Mallya A, Lazebnik S. PackNet: adding multiple tasks to a single network by iterative pruning. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition; 2018. p. 7765–7773.

[19] Van de Ven GM, Tolias AS. Three scenarios for continual learning. arXiv preprint arXiv:1904.07734. 2019.

[20] Hegde N, Muralidhara S, Schuster R, Stricker D. Modality-incremental learning with disjoint relevance mapping networks for image-based semantic segmentation. In: Proceedings of the 2025 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV); 2025 Feb. p. 5540–5549.

[21] Barbato F, Camuffo E, Milani S, Zanuttigh P. Continual road-scene semantic segmentation via feature-aligned symmetric multi-modal network. In: Proceedings of the 2024 IEEE International Conference on Image Processing (ICIP); 2024 Oct. p. 722–728.

[22] Pian W, Deng S, Mo S, Liu M, Guo Y, Tian Y. Modality-inconsistent continual learning of multimodal large language models. arXiv preprint arXiv:2412.13050. 2024.

[23] Qi CR, Yi L, Su H, Guibas LJ. PointNet++: deep hierarchical feature learning on point sets in a metric space. In: Advances in Neural Information Processing Systems. 2017;30.