BioSLAM: A Bioinspired Lifelong Memory System for General Place Recognition

Peng Yin, Member, IEEE, Abulikemu Abuduweili, Shiqi Zhao, Lingyun Xu, Changliu Liu, Member, IEEE, and Sebastian Scherer, Senior Member, IEEE

Abstract—We present BioSLAM, a lifelong (lifelong simultaneous localization and mapping) SLAM framework for learning various new appearances incrementally and maintaining accurate place recognition for previously visited areas. Unlike humans, artificial neural networks suffer from catastrophic forgetting and may forget the previously visited areas when trained with new arrivals. For humans, researchers discover that there exists a memory replay mechanism in the brain to keep the neuron active for previous events. Inspired by this discovery, BioSLAM designs a gated generative replay to control the robot’s learning behavior based on the feedback rewards. Specifically, BioSLAM provides a novel dual-memory mechanism for the maintenance of: 1) a dynamic memory to efficiently learn new observations; and 2) a static memory to balance new–old knowledge. When the agent is encountered with different appearances under new domains, the complete processing pipeline can help to incrementally update the place recognition ability, robust to the increasing complexity of long-term place recognition. We demonstrate BioSLAM in three incremental SLAM scenarios as follows. 1) A 120 km city-scale trajectories with LiDAR-based inputs. 2) A multivisited 4.5 km campus-scale trajectories with LiDAR-vision inputs. 3) An official Oxford dataset with 10 km visual inputs under different environmental conditions. We show that BioSLAM can incrementally update the agent’s place recognition ability and outperform the state-of-the-art incremental approach, generative replay, by 24% in terms of place recognition accuracy. To the best of our knowledge, BioSLAM is the first memory-enhanced lifelong SLAM system to help incremental place recognition in long-term navigation tasks.

Index Terms—Continuous localization, incremental place recognition (PR), lifelong simultaneous localization and mapping (SLAM).

I. INTRODUCTION

An essential capability for long-term robotics autonomy in the open world without human assistance is lifelong simultaneous localization and mapping (SLAM) [1]. In the context of lifelong SLAM, the system needs to consider work in long-term navigation in large-scale environments and diverse environmental conditions, as depicted in Fig. 1. Current SLAM methods are mainly conducted under single-type environments, where the environmental conditions (such as illuminations, weather, and seasons) are consistent. Recent works attempt to relax the single-type assumption to accommodate diverse environments by leveraging domain adaptation techniques [2], [3] into model learning with deep neural networks. However, the domain knowledge under new scenarios can affect the localization accuracy in previous learned areas, an effect known as “catastrophic forgetting.”

In real-world long-term navigation [4], the robot may encounter complicated 3-D environments, such as campus areas, open streets, residential blocks, and commercial buildings, and each place has its unique patterns in place recognition (PR). The robot platform cannot collect datasets under all scenarios at once and train the localization module in a supervised manner. A naive solution for incremental observations is to source additional data for model adaptation with a new scenario; however, this adaptation is not feasible when the goal is to ensure the uninterrupted and long-term operation of the robot since it causes catastrophic forgetting of previous knowledge. Moreover, changes in environments can be sudden, e.g., rapid illumination and weather changes, while it may take too long for traditional learning-based approaches to react to the changes. Given the abovementioned consideration, the main challenges for lifelong PR include the following.

1) Various Environmental Conditions: the appearances of the same area under different environmental conditions will be represented with different patterns.
2) Diverse Scenarios: the robot platform will encounter different 3-D environments in large-scale navigation tasks, and most areas are a combination of different types.
3) Nonstop Training: the robot will accumulate new datasets, and model fine-tuning is usually required to improve localization performance for new scenarios.

Although the topic of long-term SLAM [5], [6], [7], [8], [9] has been well studied in the past decade, in this work, we narrow down the scope to the lifelong PR in long-term navigation and
proposed BioSLAM, that can continuously relocalize to new environments without sacrificing recognizing ability in previously seen environments. In our previous work [10], we notice that cross-domain appearance differences will significantly affect the recognizing performance; the recognizing module encounters the catastrophic forgetting problem, where it is only robust to the most recently trained scenarios. In contrast, humans and animals do not suffer from catastrophic forgetting, and short-term and long-term memory mechanisms exist within the hippocampus [11] and the front lobe of the brain [12], which plays the main role in lifelong knowledge updating. Recently, new evidence from fMRI studies in humans [13] finds that the hippocampus may “act as a librarian to retrieve the cortical books of memory,” i.e., the hippocampus can index the memories for fast retrievals. Inspired by the biological mechanism, we design two memory zones for BioSLAM, namely static memory zones (SMZs) for historical memory encoding with low frequency and dynamic memory zone (DMZ) for quickly memory reply, and propose a dual-memory selection mechanism to balance the short-term adaptation for new observations and long-term memory retention for historical knowledge. Specifically, BioSLAM also develops a sleeping cycle for memory consolidation within SMZ, which is also inspired by a similar mechanism in the hippocampus [14]. Based on the abovementioned mechanism, BioSLAM has the ability to achieve long-term PR.

The evaluation methods [15] for the traditional PR using supervised learning approaches do not apply to lifelong systems. The adaptation capability reflects the performance of lifelong systems concerning new observations and the long-term memory retention of previously visited areas. In this work, we formulate two metrics, namely adaptation efficiency (AE) analysis, and retention ability (RA) analysis, and perform an extensive evaluation using three long-term datasets as follows.

1) \textit{ALITA Urban dataset}, which is focused on changing geometric patterns.
2) \textit{ALITA Campus dataset}, which is focused on changing illumination patterns.
3) \textit{Oxford RobotCar dataset}, which is an official long-term datasets with different environmental conditions.

The contributions of this article are as follows.
1) BioSLAM provides a systematic framework to learn about ever-changing environments without interruption. This framework enables incremental place feature learning for long-term autonomy.
2) Within BioSLAM, we develop a rewarding mechanism with a dynamic and SMZ, which contains the task-oriented external reward and curiosity-oriented internal reward, which can quickly adapt new patterns and maintain memorization for long-term memory retention.

The rest of this article will introduce the related works for PR and lifelong incremental learning in Section II. Section III gives the structural overview of BioSLAM. Section IV and Section V explain the details of the general place feature learning and bioinspired lifelong memory (BiLM), respectively. The experiment setup and qualitative/quantitative analysis are given in Section VI and Section VII.

II. RELATED WORKS

There are two essential modules in lifelong navigation: 1) navigation and 2) lifelong learning. The navigation task usually contains the PR or loop closure detection (LCD) module as stated in [16], [17], [18], which mainly serves as the data association for large-scale relocalization and map optimization in SLAM tasks. Lifelong learning, also known as continual, incremental, or sequential learning, aims at incrementally building up knowledge from a sequential data stream [19], [20], which is essential for long-term localization where robots will encounter many infinite environments. In the following sections, we will mainly introduce the related works in visual/LiDAR navigation and recent lifelong learning works from a robotics perspective.

A. Long-Term Navigation

In long-term navigation, the PR targets identifying the exact areas under different perspectives and environmental conditions [16].

The traditional geometry descriptors (e.g., scale-invariant feature transform [21] and oriented FAST and rotated BRIEF [22]) are widely used in visual PR because of their invariant properties to scale, orientation, and illumination changes. Based on these handcrafted features, FAB-MAP [23] build a bag-of-visual-words (BoW) architecture to achieve large-scale visual relocalization [8], [24]. iBoW-LCD [25] uses an incremental BoW scheme based on binary descriptors to retrieve matched images more efficiently. Shan An et al. [26] introduced FILD++, an incremental LCD approach via constructing a hierarchical small-world graph. With the booming of deep learning, new convolutional neural network (CNN) features provide significant improvements in feature/semantic extraction. NetVLAD [27]
combined the CNN features and a differentiable VLAD layer to enable deep learning for visual PR; and based on NetVLAD, recent deep learning approaches further improve the recognition accuracy with different networks.

Despite the success of existing PR methods, the nonlearning-based approaches are sensitive to parameter tuning under different scenarios; and learning-based techniques are trained in a supervised learning manner, restricting their generalization ability within the offline training datasets. However, in real-world and long-term tasks, the data stream is infinite with the combination of different areas under varying environmental conditions; meanwhile, robotic systems cannot stop and wait for the network model to update for newly encountered scenarios.

Except for the abovementioned place descriptor-based SLAM system, there are other remarkable works for long-term navigation. The authors in [5] provided the experience-based approach for ever-changing environments, and the robot can switch between different experience traces while maintaining the robustness of the localization system. Recently, [33] extended the experience-based long-term navigation for the UGV routine-following task, and [34], [35] further extended it to the teach-and-repeat visual navigation task for UAV systems. The authors in [36] developed the linear regression-based supervised change prediction mechanism to handle the predictable changes in the long-term navigation task. In [6], the author provided a map summarizing framework for lifelong visual navigation, where the multiple visited maps are incorporated into a joint map which shows better generalization ability for changing environments. On the other hand, dynamic changes within the 3-D environments may also cause localization failures in the long-term navigation task. The authors in [37] provided a frequency-enhanced map monitor mechanism, which can detect the regular changeable appearance in the long-term navigation task. Besides the experience traces, [38] also provides a lifelong navigation approach based on a particle filter with a hidden Markov model; this method can also long-term UGV localization over the parking areas under ever-changing parking spaces. The authors in [39] constructed graph pruning for lifelong SLAM, which can balance the graph size and mapping performance for long-term navigation requirements.

In this work, we target lifelong localization, where the place observations will be viewed only once in the sequential order. Instead of focusing on short-term localization or fixed pattern localization in most exciting PR methods or using the experience-based/frequency-based mechanisms to keep the long-term robust appearance features, we focus on how to provide the lifelong training procedure for the learnable place descriptor.

B. Lifelong Learning for Robotics

Lifelong learning, also known as continual learning, aims at providing incrementally updated knowledge in ever-changing environments. Although this area has been studied for a long time, most approaches are still restricted to simulation or toy datasets and cannot be applied in real robotic applications. As mentioned in [20], the fundamental challenge for lifelong learning is not necessarily finding solutions that work in the real world but rather finding stable algorithms that can learn in the real world and overcome the catastrophic forgetting problem. Recent works can be roughly divided into four families: dynamic architectures, regularization-based, rehearsal, and generative replay (GR) approach.

Dynamic architecture-based methods either 1) add additional parameters to the models, such as LwF [43], which use shared early feature extraction layers and fixed task layers; or 2) use model adaptation to avoid catastrophic forgetting, such as PackNet [44], which defines the mask layer to protect weights when learning new tasks. Regularization-based methods in the context of lifelong learning can add constraints to avoid overfitting to new tasks and keep inference ability for the previous mission, such as elastic weight consolidation [45] and synaptic intelligence (SI) [46]. Airloop [47] proposed a lifelong learning method for visual LCD utilizing a Euclidean-distance knowledge distillation loss on images. InCloud [48] proposed an angular distillation loss to encourage the network to preserve the structure of its embedding space. However, the abovementioned methods must deal with specific network structures and can quickly converge to undesired local optima for complex tasks. Rehearsal-based methods, on the other hand, use memory replays to enhance the knowledge from the previous tasks or processes, such as iCaRL [49], GEM [50], which use a small subset of the previous dataset to balance the knowledge distribution for different tasks. Instead of maintaining the knowledge based on past data samples, GR [51] combines the actual raw data and generated artificial data for model updating. In [52], the authors used a dual teacher-student GR method for incremental learning, where the teacher network is frozen to guide new networks, and the networks will switch roles when the student network surpasses the teacher. Meanwhile, for the memory mechanism in lifelong navigation, [53] introduced a long-short-mechanism to transfer the new observation to the long-term memory while maintaining robust localization performance under changing environments.

In the long-term localization task, the robots will encounter many areas under different viewpoints and environmental conditions. Hence, the ideal approach would be tackling the real-world localization problem in an embodied platform: an autonomous agent that can efficiently and incrementally update its localization ability with limited computation resources. Similar to [53], BioSLAM also utilizes the dual-memory mechanism to adapt to new observations while maintaining memorization ability for long-term knowledge; the significant difference is that BioSLAM introduced the rewarding mechanism (internal and external rewards) and the memory consolidation for the dual memory system.

III. PROBLEM FORMULATION AND SYSTEM OVERVIEW

This article presents BioSLAM, an incremental PR method that enables continual learning of place feature extraction modules across various domains. Various domains may encompass different weather conditions, road areas, or input signal modalities. The approach consists of two fundamental modules as follows. 1) A general place feature extraction module for encoding place features under different domains. 2) A BiLM
Fig. 2. Lifelong localization system framework. The lifelong localization system contains the following modules. 1) General place feature extraction network, which extracts the place feature from different domains. 2) BiLM system can provide short-term and long-term assistance to capture new knowledge and maintain old knowledge.

system for continual learning PR from different domains. In this section, we will first formulate the lifelong localization problem, then introduce the two key modules in the BioSLAM system.

A. Problem Formulation

We define a sequence of place observations under domain $D$ (i.e., visual or LiDAR) as $O^D = \{O_1, \ldots, O_M\}$, and a query of observations under the same domain as $Q^D = \{Q_1, \ldots, Q_N\}$.

The task of traditional PR is to learn a feature extraction function $F$ with parameter $\theta$ to help each frame in $Q^D$ find positive (or neighbor) places from the reference sets $O^D$. Let $d(\cdot, \cdot)$ denote the difference matrix (i.e., Euclidean distance). The objective is to make the feature differences of positive (or neighbor) places smaller than negative (or far-away) places by feature extraction function $F_{\theta^*}$.

$$L(Q_k^D) = d(F_{\theta^*}(Q_k^D), F_{\theta^*}(O_1^\infty)) - d(F_{\theta^*}(Q_k^D), F_{\theta^*}(O_\neq^\infty))$$

$\theta^* = \arg \min_{\theta} \sum_{k=1}^{N} L(Q_k^D)$ (1)

where $Q_k^D$ is the current $k$th query, $O_\infty^D$ is the positive reference in a neighbor range near $Q_k^D$, and $O_\neq^D$ is the negative reference away from $Q_k^D$.

In lifelong localization, the environmental domains ($D_1, \ldots, D_t, \ldots, D_T$) can vary under different environmental conditions or sensor modalities (e.g., incrementally learning from different illuminations or weathers). Thus, both references set $O^D_t$ and the query set $Q^{D_t}$ are obtained incrementally. As depicted in Fig. 2, the lifelong localization problem is to incrementally learn and update the feature extraction function $F_{\theta_t}$, that can quickly adapt its feature extraction ability in the newest domain $\{O^{D_T}, Q^{D_T}\}$, and also in parallel maintain the feature distinguish ability for previous domains $\{O^{D_t}, Q^{D_t}\}_{t=1,\ldots,T-1}$. Since we are considering lifelong learning [19], raw data of different domains are fed sequentially for one-time usage and cannot be stored for offline training. Thus, when optimizing feature extraction function $F_{\theta_t}$ at the current domain $\{O^{D_T}, Q^{D_T}\}$, we cannot access the previous raw data from $\{O^{D_t}, Q^{D_t}\}_{t=1,\ldots,T-1}$.

For time step $T$, lifelong localization can be formulated as

$$L^t(Q_k^{D_t}) = d(F_{\theta}(Q_k^{D_t}), F_{\theta}(O_1^{D_t})) - d(F_{\theta}(Q_k^{D_t}), F_{\theta}(O_\neq^{D_t}))$$

$$\theta_T = \arg \min_{\theta} \sum_{t=1}^{T} \sum_{k=1}^{N} L^t(Q_k^{D_t})$$ (2)

B. General Place Feature Extraction

For the lifelong purpose of long-term localization, we developed a general place feature extraction or general place learning (GPL) network based on our previous works in visual [10] and LiDAR-based [54] localization, which can be referred to as the feature extraction function $F_{\theta}$ as in (2). We use the shared spherical convolution network to simultaneously achieve LiDAR and visual place localization. The spherical harmonic-based convolution can help the learned descriptor have the viewpoint-invariant propriety for the same PR. The major difference between the current GPL and our previous works is that GPL does not contain any domain-transfer module, which has been used to reduce the feature differences for the same areas under different domains [10]. This modification is because we want to evaluate the adaptation ability for the same network. On the other hand, no task-specific network layers are used in the dynamic architecture-based lifelong modules, as stated in Section II-B. We want to avoid uncertain parameters and only focus on how memory mechanisms can help incremental learning for real-world applications.

C. Bioinspired Lifelong Memory

Inspired by the memory system in human being and other mammal animals [13], we propose a dual-memory, DMZ, and SMZ, enhanced lifelong learning mechanism to deal with catastrophic forgetting in lifelong localization. As studied in [55], to create long-term memories in our brain, we have so-called sleep circle during our sleep:
Fig. 3. BioSLAM network structure. The structure of BioSLAM includes the GPL network, the rewarding mechanism to guide the memory storing and consolidation, and the dual-memory module with SMZ/DMZ. The whole procedure includes the following. 1) New observations \( q_k \) are fed into the networks only one time and sequentially. 2) In GPL, the memory encoder converts the inputs \( q_k \) to encoded memory \( z_k \), followed by spherical convolution and VLAD layer to generate place feature \( F_k \). 3) Rewarding mechanism will estimate the external reward \( R_{ex} \) and internal rewards \( R_{in} \) to guide the memory operations. 4) Dual-memory will conduct memory storing/consolidation and replay the important (high-rewarded) memories. 5) Memory decoder generates the synthetic samples from the replayed memories. 6) Updating the GPL network module with both raw (real) samples and synthetic samples.

1) the brain can encoding our daily observation into the hippocampus zone with decay along the time;
2) a consolidation mechanism is triggered between the hippocampus and the neocortex to store essential memory traces and forget the rest traces;
3) then humans can retrieve the relative memory traces based on the consolidated ones in the neocortex.

In BioSLAM, we re-build the “sleep circle” for the lifelong localization task. As we can see in Fig. 2, the memory system of BioSLAM also includes the place feature “encoding” procedure for new observations, the memory “consolidation” controlled by a rewarding mechanism to filter out necessary traces for more extended storage, and the “retrieved” memory to re-enhance the long-term place recognition ability. Based on the abovemention architecture, the BioSLAM system constructs two major modules, the GPL module and the BiLM module, which will be investigated in Sections IV and V.

IV. GENERAL PLACE LEARNING

As shown in Fig. 3, the GPL (blue dashed box) system mainly contains two submodules: a place memory encoding module (upper part of the blue dashed box) and a generative memory reply module (lower part). All samples under different domains are fed into the system sequentially once during the lifelong learning procedure. The GPL system uses symmetric encoder-decoder networks to encode new observations to “memory codes” and decode “memory codes” into the synthetic observations. In this section, we introduce the design of the encoder, the decoder, and the place feature learning within the GPL system.

A. Place Memory Encoding and Place Feature Extraction

GPL applies the encoder module \( \mathcal{E} \) to convert raw sensor observations into the “memory codes” with VGG-based [56] networks, and the memory codes are basic materials in the BiLM system. Viewpoint differences and environmental appearance changes in lifelong localization will affect the final localization performance in real-world applications. Based on the orientation-equivalent property of spherical harmonics, we utilize the spherical convolution [10], [54] on top of the encoder module to provide viewpoint-invariant feature (also called place descriptor) to reduce the viewpoint differences in long-term relocalization.

The GPL system encodes panorama camera and 3-D local point cloud with the same encoding network structure \( \tilde{\mathcal{E}} \). For the visual inputs, we convert the raw image to \([H \times W]\) spherical perspectives. For the LiDAR inputs, instead of a single scan, we generate dense local 3-D maps using the similar voxel mapping mechanism in our previous work [57] and map the points onto the spherical projections, which have the same omnidirectional view as a panorama camera. Then, the preprocessed (visual or LiDAR) inputs \( q_k \) are fed into the encoder module for generating memory codes and place features. The memory code \( z_k \) are encoded from \( q_k \)

\[
z_k = \mathcal{E}(q_k).
\]

To extract the viewpoint-invariant place feature from \( z_k \), we utilize the spherical convolution based on the spherical harmonics [58]. In theory, spherical convolution can avoid space-varying distortions in Euclidean space by convolving spherical signals in the harmonic domain. Let \( f \) is the signal on spherical harmonic, which satisfy the viewpoint-equivariant [59] property with the signal \( \mathcal{E} \)

\[
[f \ast_{\text{SO}(3)} [H_R \mathcal{E}]](q_k) = [H_R [f \ast_{\text{SO}(3)} \mathcal{E}]](q_k)
\]

where \( H_R (R \in \text{SO}(3)) \) is the rotation operator for spherical signals. \( f \ast_{\text{SO}(3)} \mathcal{E} \) denotes the spherical convolution between \( f \) and \( \mathcal{E} \). Practically, the spherical convolution is computed...
we have a set of potential positives (close-by samples) \( \{ o^\text{pos}_k \} \) and the set of negatives (far away samples) \( \{ o^\text{neg}_k \} \). The localization loss metric is defined by

\[
L_{\text{loc}}(q_k) = \max_{i,j} \left( \| F(q_k) - F(o^\text{pos}_k) \|^2 
+ \alpha - \| F(q_k) - F(o^\text{neg}_k) \|^2, 0 \right)
\]

(7)

\[
L_{\text{rec}} = E_{q_k \sim P_{\text{data}}} \left[ L_{\text{loc}}(q_k) \right] + E_{z' \sim P_z} \left[ L_{\text{loc}}(G(z')) \right]
\]

(8)

the abovementioned equation is the triplet loss version of (1), where \( L_{\text{loc}}(q_k) \) is the localization loss metric for single query \( q_k \), \( \alpha > 0 \) is a margin to control the feature difference threshold, and \( L_{\text{loc}} \) is the localization loss for the joint \( \{ P_{\text{data}}, P_z \} \) sets.

To keep the consistency of memory encoding–decoding, we further use a reconstruction loss between encoder and decoder modules. For the encoded memory code \( z' \) and the decoded memory \( E(G(z')) \), we have

\[
L_{\text{joint}} = L_{\text{loc}} + L_{\text{rec}} + L_{\text{gan}}(G, D).
\]

(10)

The major difference between our work and the traditional GR [61] is that BioSLAM can manage the retrieved memory based on their long-term behavior instead of treating all data on the same manifold distribution. The next section will deeply investigate the lifelong memory system.

V. BIOINSPIRED LIFELONG MEMORY

In our BioSLAM, as shown in the greed dashed box of Fig. 3, the BiLM system mainly contains two modules as follows. 1) **Rewarding mechanism** to measure the memory codes’ importance (or reward calculation), the importance score will be used to manage memory consolidation and selection in the dual-memory module. 2) **Dual-memory** module to cooperate with the rewarding mechanism for long-/short-term memory storage and importance-retrieval with limited space usage; which includes SMZ and DMZ.

A. Rewarding Mechanism

When the memory system encounters a new place “memory codes” \( z \), we define the hybrid reward to control the learning behavior: an external reward \( R_{\text{ext}} \), which indicates localization ability, and the internal reward \( R_{\text{int}} \), which presents the intrinsic familiarity on observations.

1) **External Reward:** The external reward is related to the learning difficulty of data samples, which indicates their distinguishing ability in the PR task. In the standard learning paradigm, all samples with different difficulties are equally considered during the model optimization. However, humans and animals usually spend more energy and time to learn harder concepts. Inspired by animal training [63] and curriculum learning [64], it is practical to differentiate the data samples into different difficulty levels, such as “easy,” “medium,” and “hard.” For lifelong localization, the feature extraction function \( F_\theta \) may
require more “energy” or iterations to learn “hard” samples. Hence, we encourage “hard” samples to have a higher chance of retraining by defining the triplet loss to measure the sample’s difficulty and distinguishability. Based on the PR loss metric \(L_{\text{loc}}\), we define the external reward for each input as follows:

\[
R_{\text{ex}}(q_k) = L_{\text{loc}}(q_k) \tag{11}
\]

which means that if the sample \(q_k\) has a higher loss than other samples, it will require more iterations in model training. “Hard” samples tend to have a high value of \(R_{\text{ex}}\), and therefore the dual-memory module selects them for memory replay with a higher probability, as described in Section V-B4.

2) Internal Reward: The internal reward is related to the robustness of the feature extraction model when given a sample. Let \(A(q_k)\) denote the data augmentation (i.e., random rotation and random translation) for sample \(q_k\). The internal reward \(R_m\) for sample \(q_k\) is defined as the cosine distance between its feature and the feature over data augmentation

\[
R_m(q_k) = 1 - \frac{\mathcal{E}(q_k) \cdot \mathcal{E}(A(q_k))}{\|\mathcal{E}(q_k)\|_2 \cdot \|\mathcal{E}(A(q_k))\|_2}. \tag{12}
\]

The internal reward \(R_m\) also indicates the network’s familiarity with the observations. In large-scale PR, similar place patterns (street view, buildings, trees) can be frequently visited with different views; the encoder \(\mathcal{E}\) has a robust representation and low internal reward of frequently visited places. Therefore, the internal reward \(R_m\) can guide the dual memory system to focus more on unfamiliar areas. The final reward for \(q_k\) can be obtained by combining the external reward with the internal reward

\[
R_k = R_{\text{ex}}(q_k) + R_m(q_k). \tag{13}
\]

Based on this rewarding mechanism, we can evaluate all the queries and obtain a set of memory trace \(m_k = (z_k, p_k, R_k)\), where \(p_k\) is the estimated location of \(q_k\) obtained from our previous relocalization system [10], [65]. The memory traces \(m_k\) are then used as the main factor in memory operations described in Section V-B.

B. Dual-Memory and Memory Operations

The memory of human beings is highly connected with long-term memory (the neocortex) and short-term memory (the hippocampus) mechanisms within our brains. BioSLAM also constructs such paired dual-memory mechanisms,

1) Static Memory (SM): \(M_S\) is similar to the long-term memory of human beings, which stores the selected memory traces by memory consolidation.

2) Dynamic Memory (DM): \(M_D\) is similar to the short-term memory of human beings, which is a quick access memory with a portion of prestored historical memory traces. DM is automatically refreshed from the SM and connected with the memory decoder.

In general, DM has a smaller buffer size (e.g., 1000), while SM has a larger buffer size (e.g., 4000). Based on the dual-memory structure, we construct two important operations for SM: memory consolidation and forgetting, and two important operations for DM: memory refreshing and memory replay.

As shown in Fig. 3, the running mechanisms of the dual-memory module can be summarized as follows. Consider an agent employing the BioSLAM algorithm, continuously navigating an environment. Upon encountering a new observation, the memory encoder encodes the observation into memory codes. Memory consolidation enables SM to store the memory code, while forgetting allows SM to discard unimportant samples and retain significant and diverse ones. Next, DM obtains selected memories from SM via memory refreshing and replays them to a memory decoder using memory replay. The memory decoder generates synthetic samples from the replayed memory codes, which BioSLAM employs to update the agent’s PR function in combination with the new real samples.

1) SM Consolidation: As stated in [66], memory consolidation is defined as a time-dependent process by which recently learned experiences are transformed into long-lasting forms to extend the long-term memory circle. In the long-term and large-scale PR, the observations may vary not only in the spatial domain (Euclidean distance) but also in the feature domain (feature distance). Furthermore, real-world navigation tasks involve an unlimited data stream. Memory consolidation is essential to abstract concise representations and guarantees memory efficiency. To provide memory consolidation for SM, we construct a feature-spatial code \(c_k = [z_k, p_k]\) for memory trace \(m_k\), which can capture both spatial and feature properties.

At time step \(T\) and observations \(q_k^T \subseteq Q^{PT}\), the obtained new memory traces \(m_k^T = (z_k, p_k, R_k)^T\) typically consist of a large number of samples. To obtain a diverse and smaller subset of memory traces (abstraction), we use K-means-based unsupervised clustering. K-means clustering partition the \(\{m_k\}^T\) into \(K\) sets \(S^T = \{S_1^T, S_2^T, \ldots, S_K^T\}\) by feature-spatial code \(c_k = [z_k, p_k]\) to minimize the following:

\[
S^T = \arg \min_{S^T} \sum_{i=1}^{K} \frac{1}{|S_i^T|} \sum_{c_i \in S_i^T} \|c_x - c_y\|^2 \tag{14}
\]

\[
\mu_i^T = \frac{1}{|S_i^T|} \sum_{c_i \in S_i^T} c_i
\]

where \(\mu_i^T\) is the cluster centroid for cluster \(S_i^T\). The memory traces within a single cluster have similar characteristics and storing all of them is redundant and memory intensive. To optimize memory usage and retrieval efficiency, downsampling is applied within each cluster \(S_i^T\)

\[
(\tilde{S}_i^T, \tilde{\mu}_i^T) = \text{downsampling}(S_i^T, \mu_i^T), \quad |\tilde{S}_i^T| < N_{\text{max}} \tag{15}
\]

where \(N_{\text{max}}\) is the (predefined) maximum number of samples in each cluster. After the sampling process, a set of new clusters \(\tilde{S}^T = \{\tilde{S}_1^T, \tilde{S}_2^T, \ldots, \tilde{S}_K^T\}\) and their corresponding centroids \(\tilde{\mu}^T = \{\tilde{\mu}_1^T, \tilde{\mu}_2^T, \ldots, \tilde{\mu}_K^T\}\) are generated from a subset of the traces \(\{m_k\}^T\).

BioSLAM can integrate the newly generated clusters \(\tilde{S}^T\) with the existing clusters from previous steps, resulting in a total of \(S^{(M_S)} = \{\tilde{S}^1, \tilde{S}^2, \ldots, \tilde{S}^T\}\) clusters in SM with corresponding centroids \(\mu^{(M_S)} = \{\tilde{\mu}^1, \tilde{\mu}^2, \ldots, \tilde{\mu}^T\}\). If the total number
### Algorithm 1: Memory Consolidation.

**Input:** Static memory \( M_S \), new memory traces \( \{m_k\} \), maximum number of clusters \( K_{max} \)

**Output:** Updated static memory

1. Construct feature-spatial codes \( \{c_k\} = \{z_k, p_k \mid m_k\} \);
2. Calculate clusters \( \{S^T_i\}_{i=1}^K \) and centroids \( \{\mu^T_i\}_{i=1}^K \) for \( \{c_k\} \) based on Eq. (14);
3. Downsample within clusters, based on Eq. (15), to generate smaller clusters \( \{\hat{S}^T_i\}_{i=1}^K \) and centroids \( \{\tilde{\mu}^T_i\}_{i=1}^K \);
4. Append new clusters \( \{\hat{S}^T_i\}_{i=1}^K \) to \( M_S \);
5. Calculate the total cluster number \( C(M_c) \) in \( M_S \);
6. If \( C(M_c) > K_{max} \) then
   7. Memory Forgetting with Algorithm 2;
8. Update static memory \( M_S \);

### Algorithm 2: Memory Forgetting.

**Input:** Static memory \( M_S \), maximum number of clusters \( K_{max} \)

1. Load clusters \( S^{(M_c)} \) and centroids \( \mu^{(M_c)} \) from static memory \( M_S \);
2. Calculate the number of forgettable clusters \( K^* = |\mu^{(M_c)}| - K_{max} \);
3. Calculate the distance matrix \( d_{(i,j)} \) between every pair of cluster centroids on Eq. (16);
4. While repeat \( K^* \) times do
   5. Find most similar cluster pairs \( (i^*, j^*) \) based on Eq. (17);
   6. Remove cluster \( i^* \) from static memory \( M_S \) and distance matrix \( d_{(i,j)} \);

of clusters \( C(M_c) = |\mu^{(M_c)}| \) exceeds the maximum threshold \( K_{max} \), similar clusters are merged to prevent memory overflow, as described in the memory forgetting mechanism Section V-B2. The consolidation mechanism is shown in Algorithm 1.

2) SM Forgetting: As mentioned in the previous section, the space allocated to SM is limited in long-term lifelong learning. Memory forgetting, a crucial operation in SM, is implemented to eliminate redundant memory clusters that are too similar to the existing ones. If the number of current clusters exceeds the maximum limit of clusters \( C(M_c) > K_{max} \), the memory forgetting mechanism removes \( K^* = C(M_c) - K_{max} \) clusters. To accomplish this, we first calculate the cluster similarity based on the distance matrix \( d_{(i,j)} \) between every pair of cluster centroids

\[
d_{(i,j)} = \|\mu_i - \mu_j\|, \forall \mu_i, \mu_j \in \mu^{(M_c)}. \tag{16}
\]

We search for cluster pairs \( (i^*, j^*) \) with the smallest distance between their centroids and remove one of the clusters \( i^* \) from each pair. This removal procedure is repeated \( K^* \) times

\[
(i^*, j^*) = \arg \min_{i,j} d_{(i,j)}. \tag{17}
\]

This method allows us to maintain a diverse set of memory clusters while removing any redundant clusters. The memory forgetting mechanism is presented in Algorithm 2.

3) DM Refreshing: DM is similar to the short-term memory of humans, brief and storage limited. To effectively replay important memory traces from DM, we need to refresh it periodically and clone relevant memory traces from SM to DM. This is done through memory refreshing mechanisms, where DM \( M_D \) obtains memory traces \( \{m_k\} \) from SM \( M_S \) through importance sampling

\[
M_D = \text{importance\_sampling}(\{m_k\}, \{w_k\})
\]

\[
m_k = (z_k, p_k, R_k) \sim M_S, \quad w_k = \gamma^n(m_k), R_k \tag{18}
\]

where importance weights \( w_k \) are determined by the reward \( R_k \) and the time-decaying factor \( \gamma^n(m_k) \). Here, \( \gamma (0 \leq \gamma \leq 1) \) is a predefined decay parameter, and \( n(m_k) \) denotes the replayed time (or revisited time) for the trace \( m_k \). Traces with higher rewards are assigned higher sampling weights as they have lower localization ability and robustness, requiring more attention from BioSLAM. Moreover, new traces are assigned higher sampling weights since the network’s ability to learn samples with many occurrences has reached an upper limit, and storing samples replayed many times is unnecessary. The decaying mechanisms also encourage DM to be more curious about new traces. These reward decay mechanisms are inspired by the decaying factor in human memory [67], which suggests that repeated learning of the same things decreases the boost in memorization.

4) DM Replay: During lifelong learning, DM replays selected memory traces to the memory decoder for generating synthetic samples, as stated in Section IV-B. Importance sampling is used to obtain replayed memories \( \{z_k'\} \) from DM \( M_D \) in the same manner as the refreshing memory mechanisms

\[
\{z_k'\} = \text{importance\_sampling}(\{z_k\}, \{w_k\})
\]

\[
m_k = (z_k, p_k, R_k) \sim M_D, \quad w_k = \gamma^n(m_k), R_k. \tag{19}
\]

We use memory decoder \( G \) to generate replayed samples \( \{\hat{q}_k\} \) from memories \( \{z_k'\} \)

\[
\hat{q}_k = G(z_k'). \tag{20}
\]

Both new observations \( \{\hat{q}_k\} \) and generated samples \( \{\hat{q}_k\} \) are used to train the GPL network by minimizing the total loss (10). The overall lifelong learning algorithm of BioSLAM is shown in Algorithm 3.

## VI. EXPERIMENT SETUP AND CRITERIA

In this section, we will introduce the experimental setup for lifelong localization. Unlike traditional localization tasks, lifelong localization requires recorded data that include either long-term differences or large-scale geometric differences. For these reasons, we built our own data collection platform and created our own lifelong localization datasets, including the ALITA Urban dataset and the ALITA Campus dataset [70]. Then, we evaluated the performance of BioSLAM on the official Oxford RobotCar dataset.
Algorithm 3: Lifelong Learning with BioSLAM.

Input: Initial place feature extraction model $F_0$ with parameters $\theta$, Initial static memory $M_s = \emptyset$ and dynamic memory $M_d = \emptyset$

1 for $T = 1, 2, \cdots$ do
2  Obtain observation set $\{q_k\}$
3  while repeat until converge do
4    Generate replayed samples $\{\hat{q}_k\}$ from dynamic memory based on Eq. (19) and (20);
5    Calculate loss $L_{\text{joint}}$ using real samples $\{q_k\}$ and replayed samples $\{\hat{q}_k\}$ based on Eq. (10);
6    Calculate gradient $\frac{dL_{\text{joint}}}{d\theta}$ then optimize $F_\theta$ with parameters $\theta$ by gradient descend;
7    Calculate rewards $\{R_k\}$ of observations $\{q_k\}$ based on Eq. (13);
8    Static memory $M_s$ consolidation based on Algorithm 1;
9    Dynamic memory $M_d$ refreshing based on Eq. (18);
end while
end for

Fig. 5. Data collection platform. The platform can record the omnidirectional visual inputs, Velodyne VLP-16 LiDAR inputs, and Xsens MTI IMU data on an Nvidia Jetson AGX Xavier. We utilize the LiDAR odometry [68] to generate the relative odometry for each trajectory and GNSS or generalized-ICP [69] to estimate the relative transformation between different trajectories.

TABLE I
Comparison Between Different Datasets

| Dataset          | Domains                        | Scales (km) |
|------------------|--------------------------------|-------------|
| ALITA Urban      | Areas: Street, residential, terrain | 120 x 1     |
| ALITA Campus     | Input modality: Lidar, visual   | 4.5 x 8     |
| Oxford RobotCar  | Weather and road conditions    | $\approx$ 10.0 |

A. Data Collection Platform

Fig. 5 shows our data collection platform, which includes an omnidirectional camera, a Velodyne VLP-16 LiDAR device, an inertial measurement unit (Xsense MTI 30, 0.5° error in roll/pitch, 1° error in yaw, 550 mW), and an embedded GPU device (Nvidia Xavier, 8 G memory). To collect time-synced LiDAR projection and omnidirectional images, we first generated dense 3-D maps through well-known LiDAR odometry [68]. Then project the point cloud within a certain distance (default is 30 m) to the spherical projections, which have the same perspective as the omnidirectional images. We will revisit the same area under large-scale and long-term assumptions in lifelong localization. To provide the relative ground truth position between different visits: to outdoor environments, we rely on the GNSS system and generalize-ICP [69] to estimate the relative transformation; for indoor environments, we mainly rely on generalize-ICP. Please note that we cannot guarantee the meter-level global absolute localization, but we can provide accurate relative localization, which is enough for the lifelong localization task. Based on the collected datasets, we have hosted a general PR Competition for long-term PR. For more details on the data collection platform and the datasets, please refer to our dataset paper [70] and competition site.

1[Online]. Available: http://gprcompetition.com/

B. Lifelong Localization Datasets and Learning Settings

We intend to analyze the lifelong learning performance under large-scale, multimodal, and long-term three perspectives. To this end, our localization datasets include three tracks as follows.

1) ALITA Urban Dataset: As shown in Fig. 6, is targeted toward large-scale lifelong learning performance. We collected 50 trajectories within the city of Pittsburgh, focusing only on LiDAR inputs within a short-term drive, as our main concern is large-scale localization. The total trajectory distance for this dataset is 110 km. 729 query and database frames are selected from 11 trajectories to construct the test set.

2) ALITA Campus Dataset: As shown in Fig. 7, is targeted toward multimodal lifelong learning performance. Both LiDAR inputs and visual inputs are collected, and we picked up ten trajectories within Carnegie Mellon University. Each trajectory is revisited eight times under different day and night-time conditions to meet the long-term requirements. The test set contains 739 query and database frames.

3) Oxford RobotCar Dataset [71]: It is a widely used public dataset aimed at long-term lifelong learning performance. This dataset covers over 1000 km of driving from May 2014 to December 2015, providing long-term observations. As our main concern is long-term localization, we use a subset of the Oxford RobotCar dataset that includes long-term behavior and different weather conditions.

We use the lifelong learning procedure depicted in Fig. 2 to feed the sequential data stream for all datasets. In our experiments, a feature extraction function $F_\theta$ continually learns from different domains $(D_1, \ldots, D_t, \cdots)$, where the domain of each dataset is different. For the ALITA Urban dataset, domains are different areas, such as commercial buildings, parks, and residential areas. For the ALITA Campus dataset, domains are different input modalities, such as LiDAR input, daylight visual input, and nighttime visual input. For the Oxford RobotCar
dataset, domains are different weather and road conditions, such as sunny days, daylight with roadworks, night, snowy days, cloudy days, and rainy days. It is worth noting that each data sample is only fed into the system once, and BioSLAM does not save a copy of that data. The comparison between different datasets is presented in Table I.

C. Performance Evaluation

During lifelong place learning, we mainly evaluate the online localization performance through the weighted recall (WR) of top-6 retrievals over lifelong learning, which is defined by $\text{WR} = \sum_{k=1}^{6} \omega_k r_k$, $\omega_1 = 0.5, \omega_k = 0.1$ for $k \neq 1$. $r_k (1 \leq k \leq 6)$ denotes Recall@k, which measures the percentage of correctly localized queries using top-k elements returned from the database. To be considered a correct match, the query image must have at least one of the top-k retrieved reference images within the predefined neighbor range (e.g., 3 m) from the ground truth position.

In lifelong PR, our focus is on the WR curve during incremental learning by continually feeding a data stream from different domains, as opposed to evaluating the WR on fixed PR models for descriptors, as is the case with classical PR. We also introduce unique matrices for evaluating the performance of lifelong PR, namely AE for measuring fast learning ability in the current new domain, and RA for measuring the ability to memorize without
forgetting the previous domains when training samples from previous domains are no longer available. When optimizing the feature extraction function \( \mathcal{F}_\theta \) in the current domain \( D_T \), the AE is defined as the WR on observations and queries from the current domain \( \{O^{D_T}, Q^{D_T}\} \). Meanwhile, the RA is defined as the WR on observations and queries from the previous domains \( \{O^{D_t}, Q^{D_t}\} | t = 1, ..., T - 1 \).

D. Baselines

As our PR task involves different sensor modalities, nonlearning/learning methods, and lifelong/nonlifelong methods, it is not feasible to cover all the relevant state-of-the-art methods. Therefore, we focus on performance comparisons from a 2-D perspective and exclude point-like 3-D methods [72]. To compare our proposed method, we select well-known nonlearning methods (BoW [73] and CoHOG [74]), learning-based methods (NetVLAD [27] and RegionVLAD [29]) and lifelong-based methods (GR [61] and SI [46]). All of the learning-based baselines and our proposed method were incrementally trained on the same sequential data during the experiments. Among the abovementioned methods, GR and SI are the most related and essential baselines to BioSLAM. Although both BioSLAM and GR use memory replay, BioSLAM has more efficient and reasonable memory replay mechanisms. This is due to two reasons: first, BioSLAM replays samples according to their reward (importance), while GR replays them randomly and evenly. Second, BioSLAM has SM, which refreshes the DM buffer to keep diverse and important memory traces, making it easier to adapt to new trajectory observations.

VII. EXPERIMENT ANALYSIS

In this section, we analyze the lifelong PR results on large-scale city areas, multimodal Campus scenarios, and long-term Oxford RobotCar datasets.

A. Large-Scale City PR

In the localization task under city-scale environments, robots may encounter various 3-D geometric structures, such as open streets, bridges, parks, and residential areas. We evaluated the performance of BioSLAM in a large-scale lifelong learning scenario using the ALITA Urban dataset. We divided the 50 trajectories covering a distance of 120 km within the city into three different areas based on their geometric properties: area 1 for commercial buildings, area 2 for parks, and area 3 for residential districts. BioSLAM and the baseline methods learn the place features incrementally by continually feeding the trajectory observations from different areas.

Fig. 8(a) illustrates the WR curve of trajectory observations within area 1, area 2, and area 3, respectively. The vertical dotted line indicates the epoch when the trajectory observations switched from one area to another, e.g., the training observations switched from area 1 to area 2 at epoch 120. Fig. 8(b) presents the average WR curve of all trajectories during training. It is evident that BioSLAM outperforms other methods during training and achieves at least a 14% improvement in terms of final average recall. When the training observations switched from area 2 to area 3 at epoch 240, BioSLAM’s performance drop on previous trajectories is much smaller than that of other methods, as shown in Fig. 8(b). This is because BioSLAM renews necessary previous knowledge by replaying related memory traces. While GR only replays randomly, BioSLAM replays essential and highly rewarded memory traces, leading to a higher convergence rate and better final performance than other baselines.

B. Multimodal Campus PR

Training PR models separately on independent modalities, such as LiDAR and visual, can be inefficient as no information is shared between them. In this study, we demonstrate the effectiveness of BioSLAM in more practical settings where the model benefits from solving PR using multiple modalities,
such as LiDAR, daylight vision, and nightlight vision. First, knowledge gained from one modality can help to better and more quickly understand other modalities since the modalities are not completely independent in PR. Second, generalization across multiple modalities may lead to the acquisition of more universal knowledge that applies to unseen modalities, a phenomenon that is also observed in infants’ learning [76], [77]. We evaluate the performance of BioSLAM in multimodal lifelong learning scenarios using the ALITA Campus dataset. Both BioSLAM and the baseline methods incrementally learn the place features from different modalities (LiDAR, daylight vision, and nightlight vision) by continually feeding trajectory observations from each modality.

Fig. 9(a) presents a performance comparison between BioSLAM and baselines on different modalities of the ALITA Campus dataset. The vertical dotted line indicates the epoch when the trajectory observations switched from one modality to another, e.g., the training observations switched from LiDAR to daylight visual image at epoch 600. During the initial 0 ∼ 600 epochs, we train the PR model on the LiDAR inputs, and all methods show performance improvement across all modalities. This result supports the idea that the knowledge gained from one modality can aid in better and faster understanding of other modalities. Between 600 ∼ 1200 epochs, we train the model on daylight visual images. The performance of all methods improves in daylight visual images, but the LiDAR performance drops around epoch 600 due to the significant difference in the LiDAR and visual images. However, BioSLAM exhibits a smaller performance drop than other methods, indicating its ability to learn new observations in a new modality without forgetting previous modalities. Between 1200 and 1800 epochs, we train the model on nightlight visual images. Here, we observe an increase in performance in nightlight visual images for all methods, and BioSLAM also shows increased performance in LiDAR, thanks to its efficient replay mechanisms. Fig. 9(b) displays the comparison of average WR between BioSLAM and the baselines on different modalities of the ALITA Campus dataset. Initially, the performance of all methods is comparable. However, as new observations from new modalities are added, BioSLAM exhibits faster and better convergence compared to the baselines. Moreover, BioSLAM outperforms the other methods by 10% in terms of final average recall.

C. Long-term PR

In long-term localization, robots need to navigate through various weather and road conditions, such as heavy rain, night, direct sunlight, snow, and also account for changes in road and building structures over time. To evaluate BioSLAM’s performance in long-term lifelong learning, we use the Oxford RobotCar dataset. We use a subset of the original Oxford RobotCar dataset that includes diverse weather and road conditions, such as sun, roadworks, night, snow, cloud, and rain. BioSLAM and the baseline methods incrementally learn the place features by continually feeding the trajectory observations captured in various weather and road conditions.

Fig. 10(a) depicts the WR curve of trajectory observations captured under various weather and road conditions. The vertical dotted line indicates the epoch when the trajectory observations switched from one condition to another, e.g., the training observations switched from sunlight vision to roadworks at epoch 60. As shown, BioSLAM addresses catastrophic forgetting by retraining necessary previous knowledge and replaying related memory traces. When the observations switch from nightlight to snowy day at epoch 120, BioSLAM’s performance drop on previous trajectories (nightlight) is significantly smaller than other methods. A similar phenomenon is observed when the observations switch from snow to cloud at epoch 160.

Fig. 10(b) shows the average WR curve of all trajectories during training. As demonstrated, BioSLAM outperforms other
methods during training and achieves at least an 8% improvement in final average recall compared to different baselines. Although both BioSLAM and GR utilize memory mechanisms to recall previous samples, BioSLAM exhibits more efficient learning due to its rewarding and memory selection mechanisms. The comparison between BioSLAM and GR in different buffer sizes (for DM) on the Oxford RobotCar dataset is shown in Fig. 11. It is evident that BioSLAM consistently outperforms GR across various buffer sizes, showcasing its superior efficiency, especially in small memory settings.

D. Adaptation, Retention and Generalization Ability

To clearly show the lifelong property of BioSLAM, in this section, we will show the AE and RA of ALITA Urban (large-scale localization) and Oxford RobotCar (long-term localization) datasets. AR is measured as the WR in the current domain. RA is measured as the WR on the previous domains.

Fig. 12 illustrates the performance of different methods on the ALITA Urban dataset in terms of AE and RA, while Fig. 13 shows the same metrics for the Oxford RobotCar dataset. The AE metric represents the fast adaptation ability to a new domain, and BioSLAM achieves the highest AE compared to the baselines on both datasets, demonstrating its fast-learning capability. On the other hand, RA represents the memorizing ability without forgetting previous domains. BioSLAM surpasses the other methods in terms of RA, thanks to its efficient rewarding and dual-memory mechanisms. In the Oxford dataset, the RA of BioSLAM almost monotonically increases, whereas other methods can exhibit a decrease in some domains (e.g., epoch 120 180 for training with nighttime observations). BioSLAM’s RA remains stable as it leverages its efficient replay mechanisms to replay important
samples from previous domains, such as sun and roadworks, to overcome catastrophic forgetting.

In addition to the lifelong learning metrics AR and RA, we also evaluate the generalization ability of the final trained model on a fixed test set, similar to classic PR tasks. Fig. 14 displays the comparison between BioSLAM and other baselines in terms of top-k recall on the test set of the ALITA Urban dataset. Despite some nonlifelong learning methods being designed or trained for offline evaluation, BioSLAM still outperforms baselines on classic (nonlifelong learning) offline test set evaluation. We then report the generalization ability of the final trained model on the fixed test set in terms of WR, as illustrated in Table II.

On the ALITA Urban dataset, BioSLAM outperforms a state-of-the-art lifelong learning method GR by 7.6%. On the ALITA Campus dataset, BioSLAM outperforms a lifelong learning method GR by 15.6%. Note that, this article primarily focuses on incremental and lifelong learning scenarios, and thus the most crucial evaluation metric is the recall curve over incremental learning (as shown in Figs. 8 and 9). While some nonlifelong learning methods (i.e., CoHOG) may perform well on recall for a fixed test set, these methods cannot learn incrementally, and therefore their performance is limited. Consequently, lifelong learning methods have a higher potential in wider and dynamic real-world environments.

E. Ablation Study

As outlined in Section V, BioSLAM stands out from previous lifelong learning methods due to several novel mechanisms, including:

1) external reward $R_{ex}$ to indicate localization performance;
2) internal reward $R_{in}$ to indicate the robustness of feature representation;
3) SM consolidation to abstract concise memory traces, where clustering (14) is a key component;
4) DM refreshing to effectively replay diverse and important memories, with the time-decay mechanism (18) for importance weight being critical.

To evaluate the effectiveness of these mechanisms, we compared BioSLAM with the following variants:

1) w/o $R_{ex}$, which does not apply external reward and leads to $R_{k} = R_{in}(q_{k})$;
2) w/o $R_{in}$, which does not apply internal reward and results in $R_{k} = R_{ex}(q_{k})$;
3) w/o consolidation-clustering, which does not use clustering in SM consolidation, resulting in Algorithm 1 directly storing all memory traces in SM;
4) w/o time-decay, which does not use the time decay factor in importance sampling, which is equivalent to setting $\gamma = 1$ in DM refresh (18).

Note that these variants follow the control variate method, covering all important mechanisms of BioSLAM without overlapping functionalities.

Fig. 15 presents the results of the ablation study conducted on the Urban dataset. It is evident that BioSLAM outperforms its variants, and the removal of any of its components leads to a significant drop in performance. To facilitate a clear comparison of the results obtained from different variants, we present the ablation study conducted on different datasets in terms of the final average recall during lifelong learning, as depicted in Table III. Notably, the larger drop in performance is observed
when removing internal rewards, highlighting the significance of the internal reward as an indicator of feature representation robustness for retrieving memories. While the performance of the “w/o cluster-consolidation” variant is close to that of BioSLAM, the latter is more memory efficient by clustering and downsampling.

F. BioSLAM Feature Property

In this section, we analyze the learned features of BioSLAM, denoted by $\mathcal{F}_b(q_k)$, using similarity matrix and principle component analysis (PCA) on ALITA Urban and Campus datasets. The similarity matrix $M_{\text{sim}}$ is computed by taking the cosine similarity between the reference $O^i_D$ and query $Q^j_D$ features, where $M_{\text{sim}}(i, j) = \cos(\mathcal{F}(O^i_D), \mathcal{F}(Q^j_D))$. A high-contrast similarity matrix indicates that the learned feature representation, $\mathcal{F}$, has a strong ability to express and discriminate between clusters, as evidenced by high similarity values for similar places and low values for dissimilar places.

The similarity matrix of BioSLAM during lifelong learning on the ALITA Urban dataset is presented in Fig. 16(a), where the left column illustrates the sampled trajectories from area 1, area 2, and area 3, respectively. The three right columns display the similarity matrices of the corresponding trajectories (from left to right) after incrementally learning from area 1, area 2, and area 3, respectively. Notably, after learning from a new area (e.g., area 2), the similarity matrix of the previous area (e.g., area 1) shows almost no decay. This observation indicates that BioSLAM still has a strong expression ability on past trajectories even when learning from different areas. The similarity matrix of BioSLAM during lifelong learning on the ALITA Campus dataset is presented in Fig. 16(b), with subfigures from left to right representing the initial step and incremental learning from LiDAR, daylight visual, and nightlight visual images. Notably, BioSLAM can differentiate not only different modalities (as seen from the three clusters from left to right) but also different trajectories within each modality. Moreover, BioSLAM can find cross-domain relationships between place observations from different modalities. For instance, the PCA results of LiDAR and night-time visual domains for trajectory 1 are relatively close and located in the lower part of the PCA visualization results.

G. BioSLAM Memory Activity

In this section, we provide a visualization of the memory buffer during lifelong learning on the Oxford RobotCar dataset. For this experiment, we used a buffer size of 1000 samples for the DM and 5000 samples for the SM. For SM, we store the samples on the hard disk since it requires a large capacity and is not accessed frequently. For DM, we store the data in RAM because it requires frequent access and does not need a large capacity.

As described in Section V-B, SM stores concise and diverse observations based on their feature and spatial properties, which are achieved through memory consolidation. For observations from a current domain, memory consolidation mechanisms cluster the observations into $K$ clusters (where $K = 20$ in our experiment) and then downsampling to keep a maximum of $N_{\text{max}}$ samples per cluster (where $N_{\text{max}} = 50$ in our experiment). Therefore, at most $K \times N_{\text{max}}$ samples were stored in the SM for a domain or trajectory ($K \times N_{\text{max}} = 10000$ in our experiment). If the total number of clusters for all domains exceeds the maximum threshold $K_{\text{max}}$ (where $K_{\text{max}} = 100$ in
Fig. 17. PCA visualization over training. (a) ALITA Urban dataset. Visualization of observations from different areas. (b) ALITA Campus dataset. Visualization of observations from different trajectories and different modalities. (a) PCA visualization on ALITA Urban dataset. (b) PCA visualization on ALITA Campus dataset.

Fig. 18. SM, DMZ, and reward ratio (normalized) of different trajectories during lifelong learning training on Oxford RobotCar dataset. (a) Number of samples in static memory. (b) Number of samples in dynamic memory. (c) Reward ratio of different domains.

our experiment) or the total number of samples exceeds the buffer size $K_{\text{max}} \times N_{\text{max}}$, the memory forgetting mechanism is triggered to reduce the number of samples in SM to keep at most $K_{\text{max}} \times N_{\text{max}}$ samples. The value of $K_{\text{max}} \times N_{\text{max}}$ is equivalent to the buffer size of the SM, which is 5000 in our experiments. Fig. 18(a) shows the number of samples in SM from different domains, such as weather and road conditions, during lifelong learning on the Oxford RobotCar dataset. In the experiment, we receive the observations from a new domain every 60 epochs. As depicted, the SM incrementally stores samples from different domains throughout the lifelong learning process. If the total number of samples in the SM exceeds the buffer size, the memory forgetting mechanism deletes some similar samples in the memory.

DM $M_d$ selects memory traces from SM with memory refreshing. Memory refreshing is based on importance sampling, where the sampling weight is proportional to its reward value. These selected memory traces are then sent to the memory decoder for training through memory replay. Fig. 18(b) illustrates the number of samples in DM from different domains during lifelong learning on the Oxford RobotCar dataset. To better understand the sample ratio between different domains, we also plot the normalized average reward for each domain in Fig. 18(c). The average reward of a domain can be computed as the average of all samples’ reward in the domain $R^{D_t} = \frac{1}{|D_t|} \sum_{q_k \in D_t} R(q_k)$. The normalized average reward of a domain is then $\hat{R}^{D_t} = \frac{R^{D_t}}{\sum_{s \in [1,T]} R^{D_s}}$. The plot shows that the DMZ holds more memory traces (samples) from domains with higher normalized average reward. Given a domain, a higher reward indicates worse performance, and BioSLAM uses higher sampling weights to retrieve more memory traces from the high-rewarded domain to achieve better performance.

H. Incremental Learning Property

To demonstrate the incremental learning property of BioSLAM, we evaluate its performance using a confidence score during lifelong learning. The confidence score is calculated as a function of the PR loss $L_{\text{loc}}$, where confidence$(q_k) = \frac{L_{\text{max}} - L_{\text{loc}}(q_k)}{L_{\text{max}}}$, and $L_{\text{max}}$ is the normalization constant set to the maximum loss value. A higher confidence score indicates a higher PR ability. We calculate the confidence score for all observations on real trajectories. Supplemental movie 1 presents the confidence score of BioSLAM during lifelong learning on the ALITA Campus dataset. Similarly, supplemental movie 2 presents the confidence score of BioSLAM during lifelong learning on the ALITA Urban dataset. As shown in the videos, the proposed BioSLAM framework enables incremental improvements in PR ability across all areas during lifelong learning.

I. Run-Time Analysis

In this section, we present the memory and time usage of BioSLAM and compare it to another method for lifelong learning. Our experiments were conducted on an Ubuntu 18.04 system, with an Nvidia RTX 2080 Ti (12 GB) GPU, Intel Core i9-7900x processors, and 64 GB memory. We report the total memory usage on ALITA Urban datasets in Table IV. As can be seen, the memory usages of BioSLAM are acceptable for current embedded system structures.

Fig. 19 displays the time usage of the BioSLAM lifelong learning procedure on the ALITA Urban dataset, where new

2[Online]. Available: https://youtu.be/eNrwUw7BWuE

3[Online]. Available: https://youtu.be/K8pS9D5lLYs
such as LiDAR, day-time, and night-time visual domains. This adaptability enables robots to achieve long-term autonomy in real-world scenarios. Supplemental movie 3 demonstrates the mechanism and performance of BioSLAM in lifelong learning. The upper part of the video illustrates the BioSLAM framework, which comprises the GPL, the rewarding mechanism, and a dual-memory module. The lower part of the video displays the confidence scores during lifelong learning.

However, BioSLAM also has some limitations for lifelong navigation. First, it cannot offer submeter level localization ability, such as traditional visual SLAM systems. This limitation arises from the triplet loss applied in place descriptor learning, which cannot support feature-level alignment for 6-D pose estimation. A potential solution could be to combine meter-level place descriptors and submeter-level features into a joint SLAM system. However, transferring both descriptors and features into the same lifelong learning framework would be another open challenge. Second, while BioSLAM enables the PR network module to learn observations incrementally under diverse domains, it cannot achieve cross-domain PR via direct transfer without learning when the target domains significantly differ (e.g., summer versus winter), as the appearance differences between such domains exceed the network’s distinguishing ability. One potential solution is to combine an experience-based approach with BioSLAM. However, determining the required number of experiences for a given area and how to fuse/delete the experience and relative place descriptors presents another significant challenge for lifelong navigation.

In this work, BioSLAM provides a memory system for lifelong PR. On the other hand, it also provides a new option for other lifelong learning tasks. Recall the network structures, as shown in Fig. 3. The functional modules related to the PR task are mainly the place descriptor extraction $F_p$ and the relative external reward $R_{ex}$ in the rewarding mechanism. For other tasks, such as 3-D segmentation, researchers can replace the place descriptor extraction network (i.e., the spherical convolution and VLAD layer) with a 3-D U-Net [78], utilize segmentation loss metric instead of triplet loss, and not need to replace the entire blocks in the lifelong memory system. However, unlike PR which can leverage self-supervised mechanisms without human labeling, providing accurate segmentation annotations will be a significant challenge. In addition, another potential option is to develop a parallel hybrid lifelong learning system for multiple tasks since the encoder module $E$ can be shared.

### VIII. Discussion and Limitations

BioSLAM provides robust and efficient lifelong learning for large-scale and long-term PR tasks. The framework adopts a dual-memory mechanism, where long-lasting memory traces are stored in the SM $M_S$, while generative memories are retrieved from DM $M_D$, enabling efficient learning of new types of observations and maintaining the lifelong memorization ability for old knowledge. This mechanism ensures robust PR under diverse conditions. In addition, BioSLAM demonstrates strong adaptability to changes in domains, as shown in the evaluation of the ALITA Campus dataset. The performance will not drop significantly when shifting from one domain to another.
and “memory replay” for continual place feature learning. The BiLM system provides a dual-memory mechanism controlled by a rewarding mechanism to guide the “memory consolidation,” “memory forgetting,” and “memory replay” to enhance the memorization of long-term traces. We investigate the large-scale and long-term PR ability in experiments with city-scale 3-D point-cloud maps, campus-scale visual-LiDAR hybrid inputs, and long-term city-scale visual inputs. Both results show that BioSLAM can significantly balance the place learning ability for new observations and maintain the memorization ability for historical observations.

Our method can be applied to low-cost mobile robots with current embedded devices, as it has a lightweight memory system that does not require saving massive streaming datasets. Another interesting direction for future work is to enable memory sharing between client agents and the cloud server. In this case, the server can be synced with data from all kinds of scenarios by various robots to update a more general PR. Finally, the BioSLAM system can be applied to other perception tasks by modifying objective functions in the rewarding mechanism according to specific requirements.

ACKNOWLEDGMENT

Any opinions, findings, conclusions, or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of ARL and NSF.

REFERENCES

[1] C. Cadena et al., “Past, present, and future of simultaneous localization and mapping: Toward the robust-perception age,” IEEE Trans. Robot., vol. 32, no. 6, pp. 1309–1332, Dec. 2016.
[2] D. DeTone, T. Malisiewicz, and A. Rabinovich, “SuperPoint: Self-supervised interest point detection and description,” in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. Workshops, 2018, pp. 337–3712.
[3] J. Stretton and P. Thompson, “Frontal lobe function in temporal lobe epilepsy,” Epilepsy Res., vol. 98, no. 1, pp. 1–13, 2012.
[4] J. G. Klinzing, N. Niethard, and J. Born, “Mechanisms of systems memory consolidation during sleep,” Nature Neurosci., vol. 22, no. 10, pp. 1598–1610, 2019.
[5] M. Zaffar et al., “VPR-bench: An open-source visual place recognition evaluation framework with quantifiable viewpoint and appearance change,” Int. J. Comput. Vis., vol. 129, no. 7, pp. 2136–2174, 2021.
[6] S. Lowry et al., “Visual place recognition: A survey,” IEEE Trans. Robot., vol. 32, no. 1, pp. 1–19, Feb. 2016.
[7] T. Barros, R. Pereira, L. Garrote, C. Premebida, and U. J. Nunes, “Place recognition survey: An update on deep learning approaches,” 2021, arXiv:2106.10458.
[8] X. Zhang, L. Wang, and Y. Su, “Visual place recognition: A survey from deep learning perspective,” Pattern Recognit., vol. 113, 2021, Art. no. 107760.
[38] G. D. Tipaldi, D. Meyer-Delius, and W. Burgard, “Lifelong localization in changing environments,” Int. J. Robot. Res., vol. 32, no. 14, pp. 1662–1678, 2013.
[39] G. Kurz, M. Holoh, and P. Biber, “Geometry-based graph pruning for lifelong SLAM,” in Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst., 2021, pp. 3313–3320.
[40] S. Chen, J. Wu, Q. Lu, Y. Wang, and Z. Lin, “Cross-scene loop-closure detection with continual learning for visual simultaneous localization and mapping,” Int. J. Adv. Robot. Syst., vol. 18, no. 5, Art. no. 172998814211050560, 2021.
[41] N. Vödisch, D. Cattaneo, W. Burgard, and A. Valada, “Continual SLAM: Beyond lifelong simultaneous localization and mapping through continual learning,” in Proc. Int. Symp. Robot. Res., 2022, pp. 19–35.
[42] H. Blum, F. Milano, R. Zurbrigge, R. Siegwart, C. Cadena, and A. Gwaler, “Self-improving semantic perception for indoor localisation,” in Proc. Conf. Robot. Learn., 2022, pp. 1211–1222.
[43] Z. Li and D. Hoiem, “Learning without forgetting,” in IEEE Trans. Pattern Anal. Mach. Intell., vol. 40, no. 12, pp. 2935–2947, Dec. 2018.
[44] A. Mallya and S. Lazebnik, “PackNet: Adding multiple tasks to a single network by iterative pruning,” in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., 2018, pp. 7765–7773.
[45] J. Kirkpatrick et al., “Overcoming catastrophic forgetting in neural networks,” Proc. Nat. Acad. Sci., vol. 114, no. 13, pp. 3521–3526, 2017.
[46] F. Zenke, B. Poole, and S. Ganguli, “Continual learning through synaptic intelligence,” in Proc. Int. Conf. Mach. Learn., 2017, pp. 3987–3995.
[47] D. Gao, C. Wang, and S. Scherer, “AirLoop: Lifelong loop closure detection,” in Proc. Int. Conf. Robot. Automat., 2022, pp. 10664–10671.
[48] J. Knights, P. Moghadam, M. Ramezani, S. Sridharan, and C. Fooke, “InCloud: Incremental learning for point cloud place recognition,” in Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst., 2022, pp. 8559–8566.
[49] S.-A. Rebuffi, A. Kolesnikov, G. Sperl, and C. H. Lampert, “iCaRL: Incremental classifier and representation learning,” in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., 2017, pp. 5533–5542.
[50] D. Lopez-Paz and M. Ranzato, “Gradient episodic memory for continual learning,” Adv. Neural Inf. Process. Syst., vol. 30, pp. 6467–6476, 2017.
[51] G. M. Van de Ven and A. S. Tolias, “Generative replay with feed-forward connections as a general strategy for continual learning,” 2018, pp. 1188–1197, Oct. 2012.
[52] D. A. Baldwin, E. M. Markman, and R. L. Melartin, “Infants’ ability to draw inferences about nonobvious object properties: Evidence from exploratory play,” Child Develop., vol. 64, no. 3, pp. 711–728, 1993.
[53] M. H. Bernstein and M. E. Arterberry, “The development of object categorization in young children: Hierarchical inclusiveness, age, perceptual attribute, and group versus individual analyses,” Devlop. Psychol., vol. 46, no. 2, 2010, Art. no. 350.
[54] O. Çiçek, A. Abdulkadir, S. S. Lienkamp, T. Brox, and O. Ronneberger, “3D U-Net: Learning dense volumetric segmentation from sparse annotation,” in Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Interv., Springer, 2016, pp. 424–432.

Peng Yin (Member, IEEE) received the bachelor’s degree in automation from the Harbin Institute of Technology, Harbin, China, in 2013, and the Ph.D. degree in mechanical engineering from the University of Chinese Academy of Sciences, Beijing, China, in 2016.

He is currently an Assistant Professor with the City University of Hong Kong, Hong Kong, China. His research interests include LiDAR SLAM, place recognition, 3-D perception, and reinforcement learning.

Dr. Yin was a Reviewer for several IEEE conferences, such as IEEE International Conference on Robotics and Automation, IEEE/RSJ International Conference on Intelligent Robots and Systems, American Control Conference, and Robotics: Science and Systems.

Abuilikemu Abudoweil received the B.S. and M.S. degrees in electrical engineering from Peking University, Beijing, China, in 2017 and 2020 respectively. He is currently working toward the Ph.D. degree in electrical engineering and robotics with Carnegie Mellon University, Pittsburgh, PA, USA.

His research interests include deep learning, robotics, and computer vision.
Shiqi Zhao received the bachelor’s degree in mechanical engineering from the Dalian University of Technology, Dalian, China, in 2018, and the master’s degree in mechanical engineering from the University of California, San Diego, San Diego, CA, USA, in 2020.

He is currently a Research Assistant with the City University of Hong Kong, Hong Kong, China. His research interests include place recognition, 3-D perception, and deep learning.

Lingyun Xu received the bachelor’s degree in computer science from the Huazhong University of Science and Technology, Wuhan, China, in 2011, and the Ph.D. degree in computer science from the University of Chinese Academy of Sciences, Beijing, China, in 2018.

She is a Research Postdoctoral with the Department of the Robotics Institute, Carnegie Mellon University, Pittsburgh, PA, USA. Her research interests include visual tracking, SLAM, place recognition, 3-D perception, and autonomous driving.

Dr. Xu was a Reviewer for several IEEE conferences, such as IEEE International Conference on Robotics and Automation, IEEE/RSJ International Conference on Intelligent Robots and Systems, IEEE/CVF Computer Vision and Pattern Recognition Conference, and International Conference on Computer Vision.

Changliu Liu (Member, IEEE) received the B.S. degree in mechanical engineering and economics from Tsinghua University, Beijing, China, in 2012, and the M.S. degree in mechanical engineering, the M.A. degree in mathematics, and the Ph.D. degree in mechanical engineering from the University of California, Berkeley, Berkeley, CA, USA, in 2014, 2016, and 2017, respectively.

She is an Assistant Professor with Robotics Institute, Carnegie Mellon University, Pittsburgh, PA, USA. Her research focuses on designing and verifying intelligent systems with applications to manufacturing and transportation.

Dr. Liu was a recipient of the NSF Career Award, Amazon Research Award, and Ford URP Award.

Sebastian Scherer (Senior Member, IEEE) received the B.S. degree in computer science from Carnegie Mellon University (CMU), Pittsburgh, PA, USA, in 2004, the M.S. and Ph.D. degrees in robotics from CMU, in 2007 and 2010, respectively.

He is an Associate Research Professor with Robotics Institute, CMU. His research focuses on enabling autonomy for unmanned rotorcraft to operate at low altitude in cluttered environments.

Dr. Scherer is a Siebel scholar and a recipient of multiple paper awards and nominations, including AIAA@Infotech 2010 and FSR 2013. His research has been covered by the national and internal press including IEEE Spectrum, the New Scientist, Wired, der Spiegel, and the WSJ.