LEARNING WHAT TO REMEMBER: REINFORCEMENT-GUIDED MEMORY FOR SAM2 IN 3D MEDICAL SEGMENTATION

Anonymous authors
Paper under double-blind review

ABSTRACT

Recent medical foundation models such as Medical SAM2 have demonstrated strong performance for 3D multi-modal segmentation by propagating information across 2D slices via a memory bank. SAM2 uses a naive first-in-first-out memory, and discards informative long-range context. As such, we empirically discover the *informative context dropping* problem: SAM2 is susceptible to dropping high-information contexts. And it can mistakenly propagate the most recent patterns instead of mining past relevant contexts. To this end, we introduce a novel *Agentic Frame Selection* (AFS) framework that that dynamically selects *important* frames based on their expected contribution to segmentation performance. In both Interactive Segmentation and In-Context Learning settings that match various real-world clinical context, AFS demonstrates state-of-the-arts performance, improving upon recent works by 4.6% in Interactive Segmentation setting and 1.8% in In-Context Learning setting.

1 INTRODUCTION

Medical foundation segmentation models (Zhu et al., 2024; Ma et al., 2025; 2024) have recently attracted significant attention due to their strong prior knowledge and promptable paradigm, enabling more generalizable segmentation.However, most existing deep learning architectures are primarily designed for 2D imagery (Wu et al., 2025; Pham et al., 2025), while medical imaging data are inherently volumetric such as CT and MRI scans (Milletarì et al., 2016; Çiçek et al., 2016). SAM2 (Ravi et al., 2024) incorporates a memory bank mechanism that propagates contextual information across contiguous slices, improving temporal and spatial coherence during segmentation. Despite its performance, SAM2 still suffers from a critical limitation in its memory attention mechanism: informative context dropping, which arises from its FIFO-based memory management strategy.

Specifically, FIFO memory banks (Ma et al., 2025; Ravi et al., 2024) are fundamentally *context-agnostic*; they discard the oldest memory entries solely based on temporal order, without considering their contextual relevance to the current slice. This design leads to two major issues. First, we identify the *informative context dropping* phenomenon, where the most informative contextual slices are prematurely removed from memory. Our experiments in Sec. 3.2 and Fig. 2 (Top) demonstrate that informative frames are discarded; yet, they are meaningful to segment the current frame. Second, SAM2 suffers from what we term the *context mis-propagation problem*. Because the model heavily biases toward the most recent observations, it tends to overfit to short-term patterns and propagate incorrect contextual cues when anatomical or lesion appearances change. As illustrated in Fig. 1, once the lesion pattern shifts, SAM2 continues relying on the most recently observed appearance rather than retrieving the most semantically relevant context from memory. Consequently, the model may misunderstand the underlying structure and produce inaccurate segmentations.

Prior works (Ding et al., 2025; Videnovic et al., 2025) propose to improve memory bank by considering motions or distracting factors. Despite improving object tracking under occluded scenarios, they still update their memory bank using FIFO, leaving informative context dropping unresolved. Developed upon SAM2, state-of-the-art ICL methods like SANSA (Cuttano et al., 2025) propose supplying support to boost the memory by learning from support samples. Nevertheless, they still

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

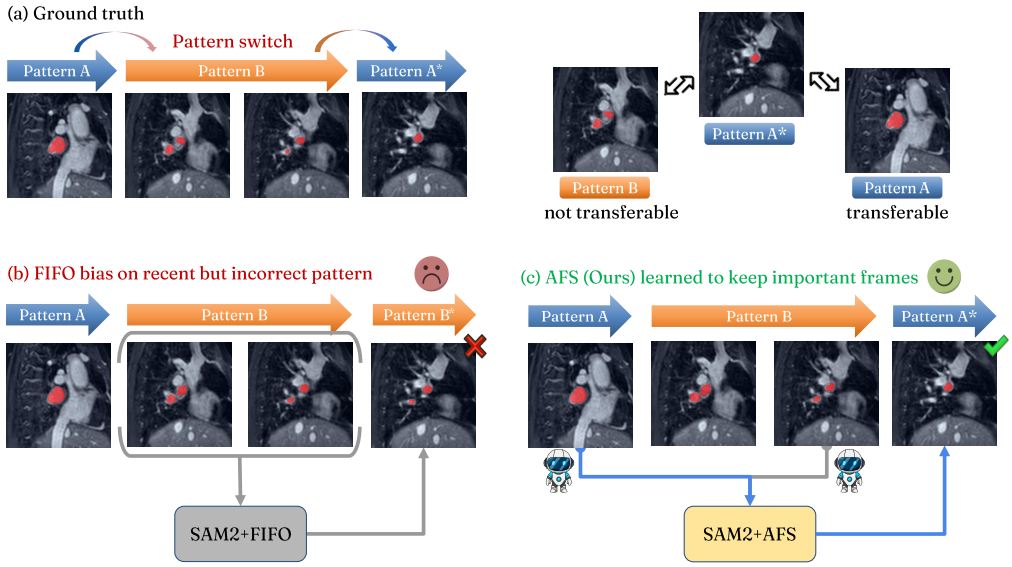


Figure 1: **Context mis-propagation phenomenon:** SAM2 employing FIFO to manage its memory bank. If there are sudden changes in contexts, e.g, from two lesion regions to one, SAM2 will be likely to mistakenly propagate the recent slice patterns to the current slice. Meanwhile, our Agentic Frame Selection dynamically learn how to retain an informative contexts relevant to the current slice, leading to more accurate segmentation.

apply naive FIFO, thus recent slice bias still persist and hamper the ability to effectively learn from the support examples.

We argue that memory bank frame selection should be inherently dynamic. Rather than relying on rigid heuristic rules, the choice of which frames to retain should adapt to both the underlying data distribution and the specific requirements of the downstream task. To this end, we propose a novel framework Agentic Frame Selection (AFS), a reinforcement learning-based memory management algorithm that learns to mine relevant visual contexts and intelligently governs both frame insertion and frame eviction within the memory bank.To demonstrate the effectiveness of AFS, we conduct experiments on both in-context learning and interactive settings, across 12 publicly available medical segmentation benchmark. Our experiment shows that our SAM2 with AFS demonstrates up to 7× lower context dropping than the original FIFO mechanism (Fig.2).

In summary, our contributions are:

- Our comprehensive experiments discover *informative context dropping* problem in prior works. SAM-2 memory drops optimal contexts, yielding beneficial gain. And it potentially propagates the biased contexts, thus mis-segmenting the current slice. Ours novel memory selection approach achieves state-of-the-art results across 11 medical datasets: +4.6% in interactive and +1.9% in ICL settings.
- We propose a novel *Agentic Frame Selection* (AFS), that adaptively selects memory slices, thus ensuring only high-quality and important slices influence future slice predictions.
- We extend our AFS to in-context and interactive settings, effectively resolve informative context dropping on both and thus demonstrating its applicability in various clinical contexts.

2 Related Works

2.1 Video Segmentation

Video segmentation has been widely studied to model temporal consistency across frames. Early approaches applied 2D models such as FCN (Lu et al., 2019) and U-Net (Azad et al., 2024), but

struggled to maintain consistency over time. SAM2 (Ravi et al., 2024) addresses this by introducing a memory bank to propagate information across frames. However, its FIFO-based design is prone to context mispropagation under dynamic changes. Subsequent works, including SAM2-Long (Ding et al., 2025), improve memory modeling with more advanced strategies, yet still face challenges in effectively adapting memory to rapidly changing contexts.

2.2 3D Medical Segmentation

3D medical image segmentation has been widely studied with models such as 3D U-Net (Agrawal et al., 2022), nnUNet (Zhou et al., 2018), and MedNeXt (Chen et al., 2024), which achieve strong performance through specialized architectures. However, these approaches are typically task-specific and require retraining for new datasets or modalities. Recent advances focus on interactive segmentation using foundation models. SAM2 (Ravi et al., 2024) extends segmentation to videos and 3D volumes via a memory bank, enabling efficient propagation across slices. Several works adapt SAM2 to medical imaging, including SAM2-Adapter (Chen et al., 2024) and DAM-4SAM (Videnovic et al., 2025). While effective, these methods rely on heuristic memory strategies such as FIFO, which suffer from memory saturation and context mispropagation in complex scenarios.

To reduce the independence of the user prompts, recent works explore in-context learning (ICL) for medical segmentation. Methods such as ICL-SAM (Hu et al., 2024), FSSAM (Xu et al., 2025), and SANSA (Cuttano et al., 2025) leverage support examples to guide segmentation. However, they still depend on the underlying memory mechanism of SAM2, limiting their ability to fully exploit contextual information. Our work addresses this limitation by solving mis-context propagation bias for both interactive and ICL settings.

2.3 Reinforcement Learning

Reinforcement Learning (RL) (Kaelbling et al., 1996) has gained significant attention in medical image segmentation, as it optimizes decision-making by enabling models to adaptively select key features. This adaptability improves segmentation performance, especially in complex tasks. Early works in this field focused on optimizing learning policies, using Deep-RL (Allioui et al., 2022) to adjust segmentation boundaries or learn to segment specific tissues, enhancing the model's ability to handle diverse medical imaging scenarios.

In video segmentation, RL has been applied to optimize memory management and prevent memory saturation by selectively retaining the most relevant frames. Techniques like Group Reward Proximal Optimization (GRPO) (Shao et al., 2024) have improved memory efficiency, while our proposed Agentic Frame Selection uses RL to choose frames that most significantly contribute to accurate segmentation.

3 Method

3.1 SAM2

SAM2 (Ravi et al., 2024) is a class-agnostic foundation segmentation model that is capable of both image and video segmentation. Users can provide visual prompts (point, box or mask) to guide it to segment specific objects.

For video segmentation, SAM2 establishes temporal connection between frames in the same video thanks to its *Memory Attention*, where image feature $\mathbf{F}_t$ of frame t is conditioned on a bank of memory features $\{E_i\}_{E_i \in \mathcal{M}}$, where $\mathcal{M}$ denotes the *memory bank*. SAM2 build memory features by fusing image features with their predicted masks. Past frames are categorized into *condition*, frames that contain visual prompts, and *non-condition*, frames in the same volume that were already predicted. SAM2 employed a FIFO mechanism for its memory bank, by selecting only memory of condition frames and temporally closest non-condition frames to influence the current frame features. When managing the memory, SAM2 makes an action: dropping the furthest and keeping the latest frame. In this research, we ask whether this strategy is always optimal. We discover the informative context dropping, where SAM2 drops *important* yet furthest frames.

3.2 INFORMATIVE CONTEXT DROPPING

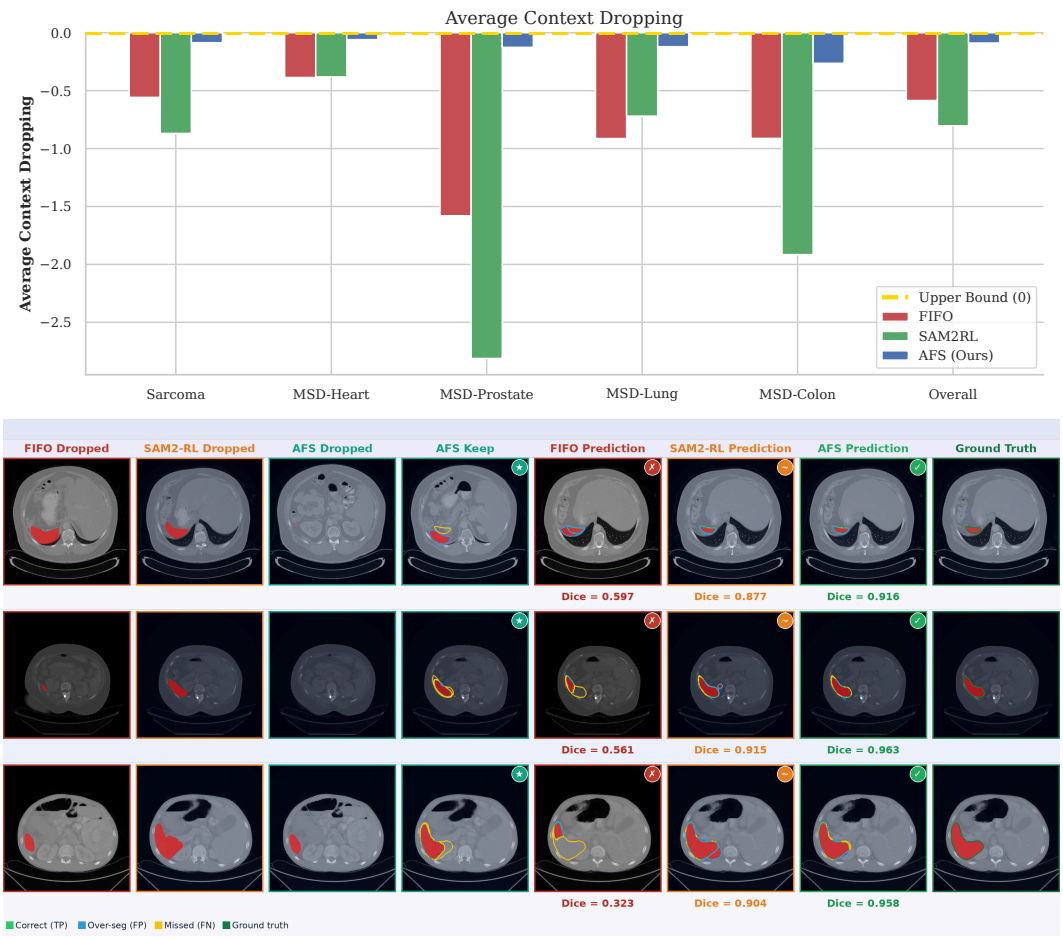


Figure 2: **Illustration of informative context dropping (Top) and context mis-propagation scenario (Bottom).** (*Top*) Quantifying the context dropping, *i.e.*, context drop, when discarding a selected frame. Quantitative results comparing Average Context Drop of FIFO, SAM2RL and our AFS, across 5 dataset and the overall. FIFO and SAM2RL show mean negative score, suggesting dropping the informative contexts. In contrast, AFS keep this score close to the upper bound, *i.e.*, near-zero optimal context loss. (*Bottom*) Qualitative comparison of mis-context propagation. FIFO and SAM2-RL discard informative historical frames, causing incorrect propagation and lower Dice scores, while AFS preserves these contexts and produces segmentations closer to the ground truth.

Both FIFO (Ravi et al., 2024) and recent SAM2RL (Adamyan et al., 2025) ignore the actual content of slices, thus prone to sub-optimal choice of action and consequently discarding informative frames relevant to the current frame, or context mis-propagation. Specifically, during inference for frame i, we have the memory bank from the previous step $\mathcal{M}_{i-1}$ and a new memory feature $\mathbf{E}_i$ that is awaiting for decision. This state induces a set of possible actions $\{a_j\}_{j\in\mathcal{M}_{i-1}\cup\{i-1\}}$, where action a_j denotes discarding frame j and taking the remaining frames as the new memory bank $\mathcal{M}_i$. Given this state and action set, each action policy π will select 1 action from this action set, denote as $\pi(\mathcal{M}_{i-1}, \mathbf{E}_i)$.

To quantitatively investigate the informative context dropping, we define the *importance* of each action as the Dice score of segmentation mask that follow the action. Then, to quantify the effectiveness of an action policy, we coin *Context Drop*, which is the difference between importance of the policy choice $\pi(\mathcal{M}_{i-1}, \mathbf{E}_i)$ and the highest importance within this action set:

$$\text{ContextDrop}(\pi) = \frac{1}{N}\sum_{i}^{N} \text{Dice}(\pi(\mathcal{M}_{i-1}, \mathbf{E}_i)) - \text{Dice}(a_i^*)$$

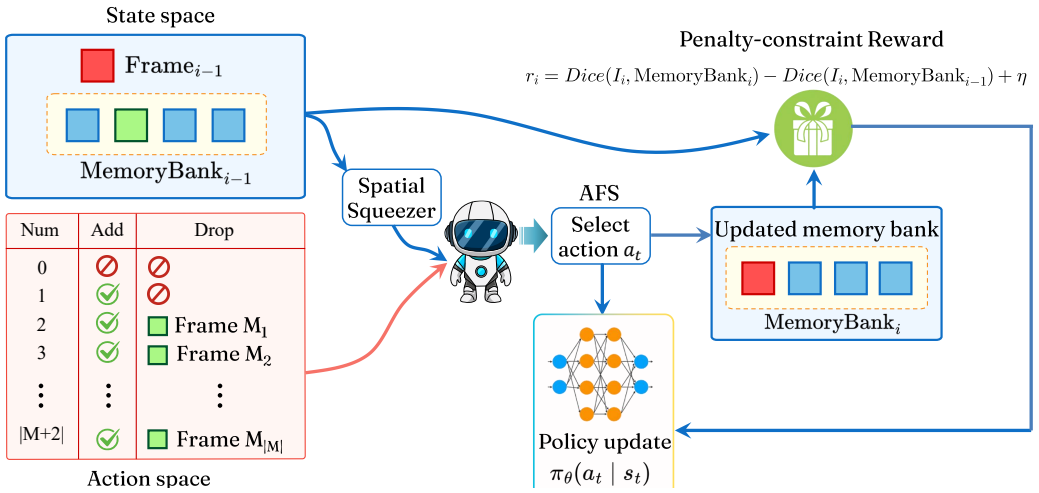


Figure 3: Agentic Frame Selection (AFS) training with reinforcement learning. Thanks to Spatial Squeezer, our agent squeezes entire volume efficiently and mine meaningful visual contexts to predict current frame.

where a_i^* denotes the most important action in the current action set. Intuitively, this metric quantifies the amount of degradation of the actual action, comparing to the optimal choice.

Two key phenomena are observed. First, FIFO and SAM2RL present highly negative Context Dropping, suggesting sub-optimal choices action selection. Second, when relevant frames are removed, the model becomes biased toward less relevant or outdated contexts, resulting in predictions that resemble temporally adjacent but mismatched frames, i.e., *context mis-propagation*. As shown in Fig. 2, SAM2 predictions drift toward incorrect regions, whereas our method remains aligned with the groundtruth.

3.3 Agentic Frame Selection

To address informative context dropping, we formalize relevant memory selection as a reinforcement learning framework. Specifically, we train a lightweight policy network to predict the most suitable actions upon the current memory bank.

To realize RL optimization for *context mining* in SAM2 memory learning, we add two components. **First**, to accommodate long sequences of 3D tokens of memory features, we introduce a Spatial Squeezer to condense long 3D sequence tokens and predict actions efficiently. **Second**, an *Action Decoder* that predict the probability of taking each actions given the squeezed spatial features.

RL Formalization. We formalize frame selection as a reinforcement learning framework. It involves an agent taking the current *state* and deciding which *action* to take, then it receives a *reward* in return. We formalize each RL component as follows. For State, we define as $\mathcal{S}$ to include the current memory features $\mathbf{E}_t$, the next image features $\mathbf{F}_{t+1}$, and all memory features in the current memory nk $\{\mathbf{E}_i\}_{\mathbf{E}_i \in \mathcal{M}}$. For Action, as described in Sec.3.2, the set of actions corresponds to the set of frames that is currently available for consideration.

Long-context Efficient Policy Network. Since spatial memory features consisting of a massive number of tokens, directly using spatial memory features leads to significant memory use. To reduce computation overhead, we design a *Spatial Squeezer* for the policy network that first squeeze spatial features, and an *Action Decoder* that predict the probability of taking each actions given the squeezed spatial features.

Spatial Squeezer module is responsible for fusing memory bank into condensed information. By employing Perceiver attention (Alayrac et al., 2022), it effectively reduces large feature maps down to a few visual tokens using learnable queries (Fig. 4). Since we include two different types of features maps in our state definition, image features and memory features, we use two different Spatial Squeezer modules for each type.

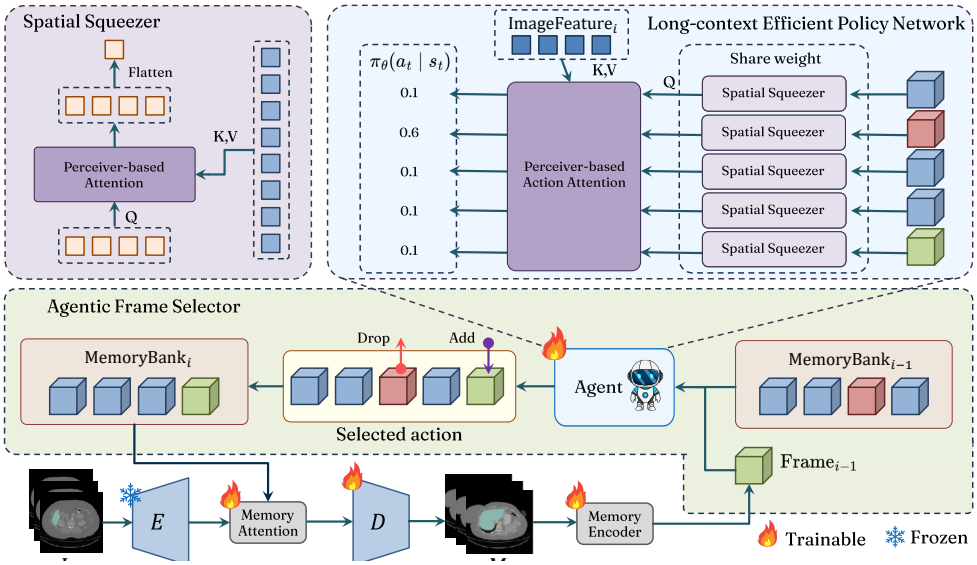


Figure 4: **Agentic Frame Selection for SAM2**. Our proposed *Agentic Frame Selection* leverages reinforcement learning to retrieve relevant contexts to optimize current frame segmentation. We formalize our Agent as a *Long-Context Efficient Policy Network*, which first squeeze 3D spatial memory features using *Spatial Squeezer*, then perform Perceiver Attention to decode the squeezed features into policy. In this figure, we omit linear projection and softmax after Action Attention for brevity.

Action Decoder also employ Perceiver attention. We use the condensed non-conditional memory features as query, and concatenate condensed image features and condensed conditional memory features to form key-value sequence.

For *Objective*, inspired by GRPO (Shao et al., 2024), while acting on the memory bank, we additionally sample a group of actions all acting on the same state and normalize the rewards within this group to obtain the advantages for Policy Gradient loss:

$$\mathbb{A}_{ij} = \frac{r_{ij} - \text{mean}(\{r_{1j}, r_{2j}, ..., r_{s_j j}\})}{\text{std}(\{r_{1j}, r_{2j}, ..., r_{s_j j}\})} \tag{1}$$

where $\mathbb{A}_{ij}$ is the advantage of sample i in group j, r_{ij} is the immediate reward for sample i in group j, s_j is the size of group j, σ is a hyper-parameter that govern the standard deviation of advantages in each group.

Finally, our policy net is optimized using the Clipped Surrogate Objective (Schulman et al., 2017), which is defined as:

$$L^{\text{CLIP}}(\theta) = \hat{\mathbb{E}}_t[\min(\frac{\pi_\theta(a_t|s_t)}{\pi_{\theta_{\text{old}}}(a_t|s_t)}\mathbb{A}_t, \text{clip}(\frac{\pi_\theta(a_t|s_t)}{\pi_{\theta_{\text{old}}}(a_t|s_t)}, 1-\epsilon, 1+\epsilon)\mathbb{A}_t)] - wH(\pi_\theta) \tag{2}$$

where $\text{clip}(\cdot, 1-\epsilon, 1+\epsilon)$ denotes clipping within range $(1-\epsilon, 1+\epsilon)$, $H(\pi_\theta)$ is entropy of the policy distribution.

Discussion: A concurrent work, SAM2RL (Adamyan et al., 2025), applies RL to SAM2-based object tracking. Yet, they do not remedy the fundamental visual context mis-propagation, inherent in SAM2's FIFO memory. This stems from two critical limitations. *First*, SAM2RL overlooks the context mis-propagation phenomenon, leaving the mechanistic reasons behind its RL gains unexplained. *Second*, its policy network is blind to visual features. To evade long-video memory overhead, it relies solely on a binary vector of frame indices. Such a vision-agnostic design is incapable of resolving visual mis-context. In contrast, our work directly addresses this core bottleneck: we empirically expose context mis-propagation through extensive evaluation, and introduce a *Spatial Squeezer* to efficiently extract grounded visual signals from long contexts without a heavy memory footprint.

As shown in Fig.2, SAM2RL still fails to keep an important frame and thus leads to sub-optimal segmentation on the current frame.

Method	**Average**		Sarcoma		Heart		Liver		Hippo		Spleen	
	DSC	IoU	DSC	IoU	DSC	IoU	DSC	IoU	DSC	IoU	DSC	IoU
SAM2-Adapter (Chen et al., 2024)	74.6	60.1	79.0	65.3	88.1	78.8	78.1	67.5	83.1	72.9	91.8	82.0
SAM2-LONG (Ding et al., 2025)	66.6	50.5	65.1	48.4	76.4	61.8	52.4	35.5	80.0	66.7	83.0	70.9
DAM-4SAM (Videnovic et al., 2025)	67.7	52.2	64.6	47.9	79.4	65.9	71.3	55.4	81.4	68.6	86.6	76.3
nnUNet (Zhou et al., 2018)	78.0	65.6	79.3	65.7	92.0	85.1	85.1	74.0	89.4	80.9	93.5	87.8
MedNeXt (Chen et al., 2024)	78.6	66.0	79.7	66.3	93.0	85.8	85.0	73.9	89.7	81.3	94.4	89.3
Ours	**79.2**	**67.0**	**80.7**	**67.6**	**93.2**	**85.8**	**85.9**	**75.6**	**90.0**	**81.7**	**94.9**	**90.3**

Method	BTCV		Prostate		Lung		Pancreas		Hepatic		Colon	
	DSC	IoU	DSC	IoU	DSC	IoU	DSC	IoU	DSC	IoU	DSC	IoU
SAM2-Adapter (Chen et al., 2024)	79.4	64.8	75.8	56.2	65.9	49.1	58.6	41.4	61.7	44.6	58.9	38.0
SAM2-LONG (Ding et al., 2025)	75.3	59.9	68.9	52.3	61.2	44.1	54.9	38.0	57.4	40.2	58.3	38.1
DAM-4SAM (Videnovic et al., 2025)	73.2	58.7	61.0	41.9	64.4	47.5	57.2	40.0	54.5	38.6	51.1	33.5
nnUNet (Zhou et al., 2018)	88.8	79.9	76.9	62.5	69.4	53.1	61.3	44.2	63.6	46.7	59.1	41.9
MedNeXt (Chen et al., 2024)	88.3	77.0	77.1	62.7	68.9	52.6	61.1	44.0	64.1	47.2	63.4	46.0
Ours	**88.8**	**80.1**	**77.9**	**63.3**	**69.9**	**53.8**	**61.9**	**44.7**	**64.2**	**47.2**	**64.4**	**47.2**

Table 1: **Performance comparison of competitive models in the interactive setting.** Average denotes the mean performance across all 11 evaluation datasets.

Method	**Average**		Sarcoma		Heart		Liver		Hippo		Spleen	
	DSC	IoU	DSC	IoU	DSC	IoU	DSC	IoU	DSC	IoU	DSC	IoU
SAM2Rad (Wahd et al., 2025)	57.7	49.8	65.4	48.6	79.2	68.5	53.8	49.0	62.4	59.4	67.4	58.3
ICL-SAM (Hu et al., 2024)	45.7	34.9	57.2	43.8	60.4	49.3	56.4	45.3	59.7	45.5	55.0	42.8
FSSAM (Xu et al., 2025)	51.2	44.3	74.7	64.2	81.6	73.1	48.5	42.4	61.9	73.1	81.4	72.8
SANSA (Cuttano et al., 2025)	66.9	57.8	72.8	62.8	84.7	77.5	68.1	**61.7**	77.6	66.7	90.1	85.8
SAM2 (Ravi et al., 2024)	68.2	59.1	75.0	65.2	85.5	77.7	62.9	56.3	77.4	66.5	90.4	86.1
Ours	**69,4**	**59.8**	**78.6**	**68.1**	**86.3**	**77.2**	**68.2**	60.9	**78.8**	**68.1**	**92.0**	**87.8**

Method	BCTV		Prostate		Lung		Pancreas		Hepatic		Colon	
	DSC	IoU	DSC	IoU	DSC	IoU	DSC	IoU	DSC	IoU	DSC	IoU
SAM2Rad (Wahd et al., 2025)	50.4	21.5	52.9	30.8	41.3	38.0	43.5	27.8	21.3	13.4	35.8	26.2
ICL-SAM (Hu et al., 2024)	45.7	33.2	61.4	47.9	40.3	31.0	26.4	17.0	24.8	15.6	15.3	12.9
FSSAM (Xu et al., 2025)	52.0	43.3	60.8	49.8	58.3	47.2	38.6	27.5	19.5	12.6	21.2	15.0
SANSA (Cuttano et al., 2025)	73.9	65.8	63.7	53.6	54.7	44.7	54.1	43.8	54.6	43.8	53.1	43.9
SAM2 (Ravi et al., 2024)	68.0	60.0	61.9	52.2	55.2	45.0	57.0	45.8	54.5	43.1	60.7	50.5
Ours	**74.4**	**66.3**	**64.6**	**53.6**	**56.6**	**46.2**	**56.0**	**45.1**	**55.3**	**43.9**	**63.7**	**52.5**

Table 2: **Performance comparison of competitive models in the 5-shot ICL setting.** Average denotes the mean performance across all 11 evaluation datasets.

Support-guided Segmentation. To demonstrate the effectiveness of our agentic-driven SAM2, we further apply our AFS module to ICL settings and achieve state-of-the-art performance. Inspired by SANSA (Cuttano et al., 2025), we directly encode support examples into the memory bank and perform segmentation on query images conditioned on the given support examples, while also employing AFS to select the most relevant memory frames to jointly supply information for segmentation.

4 Experimental results

Datasets We evaluate our method on 9 datasets from Medical Segmentation Decathlon (MSD) (Simpson et al., 2019), BTCV (Gibson et al., 2018), and Sarcoma (Vallières et al., 2015). MSD

contains 1,746 annotated 3D medical images, spanning across 10 medical segmentation tasks with various modalities, we exclude Brain Tumour dataset due to limited computational resource. BTCV comprises of 90 abdominal CT images delineating multiple organs. Sarcoma provides CT and MRI imaging data from 51 patients with corresponding Sarcoma tumour contours.

Evaluation metrics Follow prior works on segmentation, we evaluate our SAM-2 with AFS and all baselines with Dice score (DSC) and Intersection-over-Union (IoU).

Settings. We evaluate our method under two experimental settings: *ICL* and *interactive*.

In the *ICL-setting*, the model is trained in a few-shot manner, where only a limited number of annotated samples are provided during training. This setting is designed to evaluate the capability of the proposed method under low-data regimes and to simulate in-context learning scenarios.

In the *interactive setting*, the model is trained using the entire training dataset. During inference, a bounding box prompt is provided only on the first frame, and the model propagates the segmentation to subsequent frames automatically. This setting evaluates the interactive segmentation capability of the model in practical video segmentation scenarios.

Resource. All the experiments are conducted using PyTorch on 2x NVIDIA A100 80GB GPUs.

4.1 State-of-the-art comparisons

We evaluate our method under two settings: (1) the *ICL-setting*, where models are trained in a few-shot manner using only 5 annotated examples, and (2) the *interactive setting*, where models are trained on the full training set and receive a box prompt on the first frame during inference. Table 1 shows the results in the interactive setting, while Table 2 shows the results in the 5-shot ICL setting.

Interactive setting. In the interactive setting, we compare our method with recent SAM2-based approaches (SAM2-LONG, DAM-4SAM, SAM2-Adapter) and strong fully-supervised segmentation models (nnUNet and MedNeXt) on 11 medical segmentation tasks. As shown in Table 1, our method achieves the best performance on the majority of datasets, demonstrating the effectiveness of reinforcement-guided memory selection. Specifically, our method attains the highest average Dice score of **79.2%** and IoU of **67.0%**, outperforming the strongest baseline MedNeXt (78.6% Dice, 66.0% IoU). Compared to SAM2 variants, our method shows substantial improvements in both settings. In the interactive setting, we outperform SAM2-Adapter by +4.6% Dice on average (79.2% vs. 74.6%).

ICL-setting. We first compare our method with recent prompt-based and SAM-based approaches, including SAM2Rad (Wahd et al., 2025), MaskSAM (Xie et al., 2025), ICL-SAM (Hu et al., 2024), FSSAM (Xu et al., 2025), SANSA (Cuttano et al., 2025), and SAM2(Ravi et al., 2024) under the 5-shot in-context learning setting. As shown in Table 2, our method achieves the best overall performance with an average Dice score of **68.8%** and IoU of **59.4%**, outperforming the strongest prior method SAM2 (68.2% Dice, 59.1% IoU) and SANSA (66.9% Dice, 57.8% IoU).

Overall, the results from both ICL and interactive settings indicate that our reinforcement-guided memory selection effectively mitigates context mis-propagation and memory saturation in SAM2, leading to robust and consistent improvements across diverse medical segmentation tasks and data regimes.

4.2 Ablation studies.

To analyze the impact of the proposed reinforcement-guided memory selector on segmentation accuracy, memory efficiency, and contextual consistency, we compare our method with the original SAM2 (without memory selection) under varying memory bank sizes under interactive segmentation setting.

Effect of memory selection on accuracy. Our work first validates the hypothesis that reinforcement-guided memory selection improves segmentation accuracy. To this end, we compare our method with SAM2, which adopts a FIFO-based memory mechanism without selective filtering. As shown in Fig. 5, our method consistently outperforms SAM2 across all datasets. For example, on Spleen and Heart, our method achieves significantly higher Dice scores (e.g., 0.949 vs 0.831 and 0.932 vs 0.690,

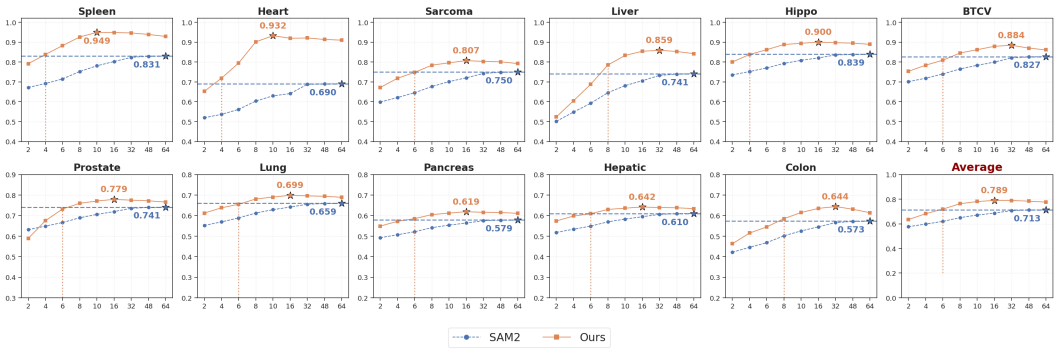


Figure 5: **Ablation study on memory selection.** We compare our method with SAM2 under varying memory bank sizes across multiple datasets. Our approach consistently achieves higher Dice scores and reaches peak performance with fewer frames. The dashed horizontal lines indicate the best performance of SAM2, showing that our method surpasses it even with a smaller memory budget.

respectively). These results confirm that selectively retaining informative frames effectively mitigates context mis-propagation and leads to more accurate segmentation.

Effect of memory selection on memory efficiency. Our work further evaluates whether the proposed memory selector improves memory efficiency.

(1) Improved performance under limited memory. This work varies the memory bank size and compares the performance between our method and SAM2. As illustrated in Fig. 5, our method consistently achieves higher Dice scores across all memory budgets. Notably, under small memory sizes (e.g., 4–8 frames), the performance gap is particularly significant, indicating that our method makes more effective use of limited memory.

(2) Faster convergence to optimal performance. It is demonstrated that our method reaches its peak performance with fewer memory frames compared to SAM2. Specifically, the performance curve of our method saturates earlier, while SAM2 requires a larger memory bank to approach its best performance. This suggests that memory redundancy in FIFO-based approaches leads to suboptimal context utilization, which is effectively alleviated by our selection strategy.

Overall, these results show that our method mitigates context mis-propagation by prioritizing semantically relevant frames, improving both segmentation accuracy and memory efficiency.

5 Conclusion

Despite achieving significant results, we based our arguments on a limited action set in order to remedy the complexity of mining key frames from the entire past. In conclusion, this work proposes a novel reinforcement learning-based memory selection framework for SAM-2, enabling more effective semantic propagation in 3D medical image segmentation. The proposed Agentic Frame Selection (AFS) method dynamically selects informative memory frames to optimize segmentation performance while mitigating context mis-propagation. Experimental results demonstrate state-of-the-art performance across multiple medical segmentation benchmarks, with particularly strong improvements on challenging sarcoma contouring tasks.

References

Alen Adamyan, Tomáš Čížek, Matej Straka, Klara Janouskova, and Martin Schmid. Sam2rl: Towards reinforcement learning memory control in segment anything model 2. *arXiv preprint arXiv:2507.08548*, 2025.

Pranjal Agrawal, Nitish Katal, and Nishtha Hooda. Segmentation and classification of brain tumor using 3d-unet deep neural networks. *International Journal of Cognitive Computing in Engineering*, 3:199–210, 2022.

Jean-Baptiste Alayrac, Jeff Donahue, Pauline Luc, Antoine Miech, Iain Barr, Yana Hasson, Karel Lenc, Arthur Mensch, Katherine Millican, Malcolm Reynolds, et al. Flamingo: a visual language model for few-shot learning. *Advances in neural information processing systems*, 35:23716–23736, 2022.

Hanane Allioui, Mazin Abed Mohammed, Narjes Benameur, Belal Al-Khateeb, Karrar Hameed Abdulkareem, Begonya Garcia-Zapirain, Robertas Damaševičius, and Rytis Maskeliūnas. A multi-agent deep reinforcement learning approach for enhancement of covid-19 ct image segmentation. *Journal of personalized medicine*, 12(2):309, 2022.

Reza Azad, Ehsan Khodapanah Aghdam, Amelie Rauland, Yiwei Jia, Atlas Haddadi Avval, Afshin Bozorgpour, Sanaz Karimijafarbigloo, Joseph Paul Cohen, Ehsan Adeli, and Dorit Merhof. Medical image segmentation review: The success of u-net. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 46(12):10076–10095, 2024.

Tianrun Chen, Ankang Lu, Lanyun Zhu, Chaotao Ding, Chunan Yu, Deyi Ji, Zejian Li, Lingyun Sun, Papa Mao, and Ying Zang. Sam2-adapter: Evaluating & adapting segment anything 2 in downstream tasks: Camouflage, shadow, medical image segmentation, and more, 2024. URL https://arxiv.org/abs/2408.04579.

Özgün Çiçek, Ahmed Abdulkadir, Soeren S Lienkamp, Thomas Brox, and Olaf Ronneberger. 3d u-net: learning dense volumetric segmentation from sparse annotation. In *International conference on medical image computing and computer-assisted intervention*, pp. 424–432. Springer, 2016.

Claudia Cuttano, Gabriele Trivigno, Giuseppe Averta, and Carlo Masone. SANSA: Unleashing the hidden semantics in SAM2 for few-shot segmentation. In *The Thirty-ninth Annual Conference on Neural Information Processing Systems*, 2025. URL https://openreview.net/forum?id=o8r3gOFTQo.

Shuangrui Ding, Rui Qian, Xiaoyi Dong, Pan Zhang, Yuhang Zang, Yuhang Cao, Yuwei Guo, Dahua Lin, and Jiaqi Wang. Sam2long: Enhancing sam 2 for long video segmentation with a training-free memory tree. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 13614–13624, 2025.

Eli Gibson, Francesco Giganti, Yipeng Hu, Ester Bonmati, Steve Bandula, Kurinchi Gurusamy, Brian Davidson, Stephen P. Pereira, Matthew J. Clarkson, and Dean C. Barratt. Multi-organ abdominal ct reference standard segmentations, February 2018. URL https://doi.org/10.5281/zenodo.1169361.

Jiesi Hu, Yang Shang, Yanwu Yang, Xutao Guo, Hanyang Peng, and Ting Ma. Icl-sam: Synergizing in-context learning model and sam in medical image segmentation. *Medical Imaging with Deep Learning*, pp. 641–656, 2024.

Leslie Pack Kaelbling, Michael L Littman, and Andrew W Moore. Reinforcement learning: A survey. *Journal of artificial intelligence research*, 4:237–285, 1996.

Yi Lu, Yaran Chen, Dongbin Zhao, and Jianxin Chen. Graph-fcn for image semantic segmentation. In *International symposium on neural networks*, pp. 97–105. Springer, 2019.

Jun Ma, Yuting He, Feifei Li, Lin Han, Chenyu You, and Bo Wang. Segment anything in medical images. *Nature Communications*, 15(1):654, Jan 2024. ISSN 2041-1723. doi: 10.1038/s41467-024-44824-z. URL https://doi.org/10.1038/s41467-024-44824-z.

Jun Ma, Zongxin Yang, Sumin Kim, Bihui Chen, Mohammed Baharoon, Adibvafa Fallahpour, Reza Asakereh, Hongwei Lyu, and Bo Wang. Medsam2: Segment anything in 3d medical images and videos. *arXiv preprint arXiv:2504.03600*, 2025.

Fausto Milletarì, Nassir Navab, and Seyed-Ahmad Ahmadi. V-net: Fully convolutional neural networks for volumetric medical image segmentation. *2016 Fourth International Conference on 3D Vision (3DV)*, pp. 565–571, 2016. URL https://api.semanticscholar.org/CorpusID:206429151.

Cuong M Pham, Phi Le Nguyen, Thanh Trung Nguyen, Vu Minh Hieu Phan, and Binh P Nguyen. Unleashing sam for few-shot medical image segmentation with dual-encoder and automated prompting. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 670–680. Springer, 2025.

Nikhila Ravi, Valentin Gabeur, Yuan-Ting Hu, Ronghang Hu, Chaitanya Ryali, Tengyu Ma, Haitham Khedr, Roman Rädle, Chloe Rolland, Laura Gustafson, et al. Sam 2: Segment anything in images and videos. *arXiv preprint arXiv:2408.00714*, 2024.

John Schulman, Filip Wolski, Prafulla Dhariwal, Alec Radford, and Oleg Klimov. Proximal policy optimization algorithms. *arXiv preprint arXiv:1707.06347*, 2017.

Zhihong Shao, Peiyi Wang, Qihao Zhu, Runxin Xu, Junxiao Song, Xiao Bi, Haowei Zhang, Mingchuan Zhang, YK Li, Yang Wu, et al. Deepseekmath: Pushing the limits of mathematical reasoning in open language models. *arXiv preprint arXiv:2402.03300*, 2024.

Amber L Simpson, Michela Antonelli, Spyridon Bakas, Michel Bilello, Keyvan Farahani, Bram Van Ginneken, Annette Kopp-Schneider, Bennett A Landman, Geert Litjens, Bjoern Menze, et al. A large annotated medical image dataset for the development and evaluation of segmentation algorithms. *arXiv preprint arXiv:1902.09063*, 2019.

Martin Vallières, Carolyn R. Freeman, Sonia R. Skamene, and Issam El Naqa. Soft-tissue-Sarcoma | A radiomics model from joint FDG-PET and MRI texture features for the prediction of lung metastases in soft-tissue sarcomas of the extremities. The Cancer Imaging Archive, 2015. URL https://doi.org/10.7937/K9/TCIA.2015.7GO2GSKS. Data collection containing FDG-PET/CT and MRI imaging data from 51 patients with soft-tissue sarcomas.

Jovana Videnovic, Alan Lukezic, and Matej Kristan. A distractor-aware memory for visual object tracking with sam2. In *Proceedings of the Computer Vision and Pattern Recognition Conference*, pp. 24255–24264, 2025.

Assefa Seyoum Wahd, Banafshe Felfeliyan, Yuyue Zhou, Shrimanti Ghosh, Adam McArthur, Jiechen Zhang, Jacob L Jaremko, and Abhilash Hareendranathan. Sam2rad: A segmentation model for medical images with learnable prompts. *Computers in Biology and Medicine*, 187:109725, 2025.

Junde Wu, Ziyue Wang, Mingxuan Hong, Wei Ji, Huazhu Fu, Yanwu Xu, Min Xu, and Yueming Jin. Medical sam adapter: Adapting segment anything model for medical image segmentation. *Medical image analysis*, 102:103547, 2025.

Bin Xie, Hao Tang, Bin Duan, Dawen Cai, Yan Yan, and Gady Agam. Masksam: Auto-prompt sam with mask classification for volumetric medical image segmentation. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 24423–24433, 2025.

Qianxiong Xu, Lanyun Zhu, Xuanyi Liu, Guosheng Lin, Cheng Long, Ziyue Li, and Rui Zhao. Unlocking the power of SAM 2 for few-shot segmentation. In *Forty-second International Conference on Machine Learning*, 2025. URL https://openreview.net/forum?id=TQtUTC3eKv.

Zongwei Zhou, Md Mahfuzur Rahman Siddiquee, Nima Tajbakhsh, and Jianming Liang. Unet++: A nested u-net architecture for medical image segmentation. *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support : 4th International Workshop, DLMIA 2018, and 8th International Workshop, ML-CDS 2018, held in conjunction with MICCAI 2018, Granada, Spain, S...*, 11045:3–11, 2018. URL https://api.semanticscholar.org/CorpusID:50786304.

Jiayuan Zhu, Abdullah Hamdi, Yunli Qi, Yueming Jin, and Junde Wu. Medical sam 2: Segment medical images as video via segment anything model 2. *arXiv preprint arXiv:2408.00874*, 2024.

A Appendix

You may include other additional sections here.