Transformer-based Imagination with Slot Attention

Yosuke Nishimoto
Graduate School of Engineering Science
Osaka University
Toyonaka, Osaka
u835711c@ecs.osaka-u.ac.jp

Takashi Matsubara
Faculty of Information Science and Technology
Hokkaido University
Sapporo, Hokkaido
matsubara@ist.hokudai.ac.jp

Abstract

World models have been proposed for improving the learning efficiency of deep reinforcement learning (RL) agents. However, it remains challenging for world models to effectively replicate environments that are high-dimensional, non-stationary, and comprising multiple objects and their interactions. We propose Transformer-based Imagination with Slot Attention (TISA), an RL agent that integrates a Transformer-based object-centric world model, policy function, and value function. The world model in TISA uses a Transformer-based architecture to handle each object's state, actions, and rewards (or costs) separately, effectively managing high-dimensional observations and preventing the combinatorial explosion of dynamics. Also, the Transformer-based policy and value functions can make decisions by considering the dynamics of individual objects and their interactions. In Safety-Gym benchmark, TISA outperforms a previous Transformer-based world model method.

1 Introduction

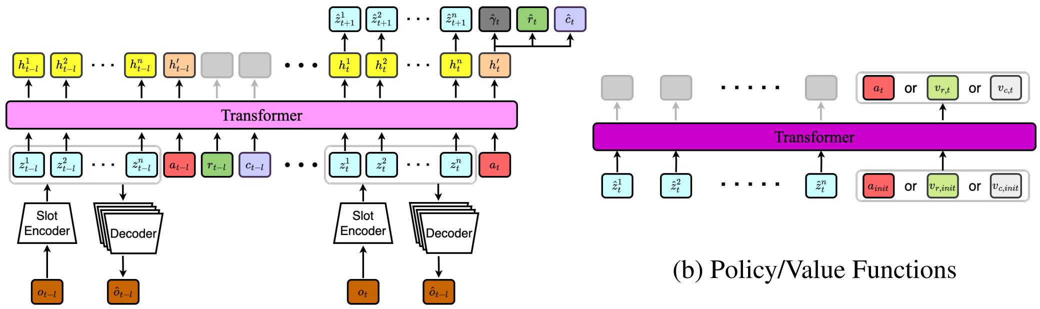
Humans understand and predict real-world dynamics through interaction with the environment [McCloskey et al., 1983]. Inspired by this mechanism, world models have been proposed for improving the learning efficiency of deep reinforcement learning (RL) agents [Ha and Schmidhuber, 2018]. These agents train the world model to replicate their observations and actions and optimize their policies within an "imagined" environment generated by the world model [Hafner et al., 2020, 2021]. Recently, deep RL agents with Transformer-based world models have achieved even higher performance [Robine et al., 2023, Micheli et al., 2023, Zhang et al., 2023]. However, it remains challenging for world models to effectively replicate environments that are high-dimensional, non-stationary, and comprising multiple objects and their interactions. In contrast, when humans are placed in such environments, they perceive the world not as a monolithic entity but by decomposing it into discrete concepts such as objects and events [Spelke and Kinzler, 2007], enabling more efficient decision-making. If integrating these cognitive mechanisms into world models, RL agents would be capable of operating more effectively even in complex environments.

Object-centric representation learning is a method that accurately represents a visual scene by segmenting it into multiple entities and extracting individual representations for each one. This approach has been applied, for example, to video prediction, where it enables the introduction of mechanisms that predict the unique dynamics of each entity and their interactions, achieving superior performance compared to conventional representation learning methods [Lin et al., 2020, Zoran et al., 2021]. Also, SlotFormer [Wu et al., 2023] demonstrated that the video prediction can be refined simply by combining object-centric representation learning with a Transformer-based autoregressive model, without the need for such specialized mechanisms. On the one hand, some studies have proposed to combine RL agents and object-centric representation learning; specifically, OCRL [Yoon et al., 2023] and EIT [Haramati et al., 2024] employ Transformers that deal with object-centric representations in both the policy and value function. This enables decision-making that focuses on

NeurIPS 2024 Workshop on Compositional Learning: Perspectives, Methods, and Paths Forward.

Table 1: Comparison of Our Proposal against Related Methods.

	SlotFormer	TWM	OCRL, EIT	TISA (proposed)
Learn object-centric representations	✓		✓	✓
Learn action-conditioned dynamics		✓		✓
Transformer-based dynamics model	✓	✓		✓
Transformer-based policy/value function			✓	✓



(a) World Model

(b) Policy/Value Functions

Figure 1: The architecture of TISA.

individual objects; however, it does not allow for decision-making explicitly based on their dynamics and interactions.

In this work, we propose Transformer-based Imagination with Slot Attention (TISA), an RL agent that integrates a Transformer-based object-centric world model, policy function, and value function. The world model in TISA employs a Transformer-based architecture and handles each object's state, the agent's action, and the reward obtained (or cost) individually. Intuitively speaking, it extracts discrete concepts in high-dimensional, complex observations, prevents the combinatorial explosion of dynamics over time, and makes the learning process more effectively. Also, the Transformer-based policy and value functions can make decisions by considering the dynamics of individual objects and their interactions. We have evaluated TISA on the Safety-Gym benchmark, which features RL environments containing various types of objects, including those directly related to rewards or costs, as well as dynamic and static objects. In this benchmark, TISA outperforms TWM [Robine et al., 2023], which is a Transformer-based world model method without object-centric representations. Table 1 presents a comparison between our proposed TISA and related methods.

2 Methods

2.1 World Model

Our world model consists of a slot-based autoencoder model and a Transformer-based dynamics model, as shown in Figure 1 (a).

Slot-based AutoEncoder Model: The slot-based autoencoder model is trained to extract object-centric representations from observations through the reconstruction. The slot encoder extracts object-centric representations from an observation o_t into n slots $(s^1, \ldots, s^n)_t$ using a Transformer called Slot Attention [Locatello et al., 2020]. The deterministic slots $(s^1, \ldots, s^n)_t$ are transformed into probabilistic latent states $(z^1, \ldots, z^n)_t$. The slots $(s^1, \ldots, s^n)_t$ are logits for categorical distributions, which are denoted by latent states $(z^1, \ldots, z^n)_t$ that each have 16 categorical variables, with each variable represented by an 8-dimensional one-hot vector [Hafner et al., 2021]. In the mixture decoder, the latent states $(z^1, \ldots, z^n)_t$ are each decoded into RGB images and (unnormalized) masks using a spatial broadcast decoder [Watters et al., 2019]. The masks are normalized across latent states $(z^1, \ldots, z^n)_t$ using a softmax function and use them as mixture weights to combine the individual RGB images into a single reconstruction $\hat{o}_t$. Figure 2 visualizes the reconstructed images for each latent state.

$$\begin{aligned} &\text{Slot Encoder:} && (z^1, \ldots, z^n)_t \sim p_\phi((z^1, \ldots, z^n)_t | o_t), \\ &\text{Mixture Decoder:} && \hat{o}_t \sim p_\phi(\hat{o}_t | (z^1, \ldots, z^n)_t). \end{aligned} \tag{1}$$

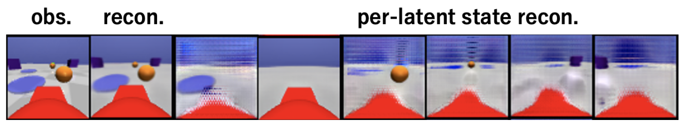


Figure 2: Visualization of (per-latent state) reconstructions.

Transformer-based Dynamics Model: The Transformer-based dynamics model predicts the future latent states $(z^1, \ldots, z^n)_{t+1}$ from the past latent states $(z^1, \ldots, z^n)_s$, action a_s, reward r_{s-1}, and cost c_{s-1} for $s = t, t-1, t-2, \ldots$. We denote the history of the latent states from time $t-l$ to t by $(z^1, \ldots, z^n)_{t-l:t}$. The Transformer-based dynamics model is composed several modules. The hidden state predictor f_ψ is a Transformer that feeds each entity of the history of the latent states $(z^1, \ldots, z^n)_{t-l:t}$, action $a_{t-l:t}$, reward $r_{t-l:t-1}$, and cost $c_{t-l:t-1}$ as tokens and predicts the deterministic hidden states $(h^1, \ldots, h^n)_t$ and h'_t as the outputs of the current latent states $(z_t, \ldots, z^n)_t$ and action a_t. In this Transformer, causal masking prevents self-attention layers from accessing future time steps in the training sequence, but the hidden states $(h^1, \ldots, h^n)_t$ can still access the current latent states $(z^1, \ldots, z^n)_t$ and action a_t. Note that the positional encoding for this Transformer is only dependent on time t but not on the index of the latent states, which takes values from 1 to n, because the prediction is expected to be equivariant to the order of the latent states. The latent states, reward, cost, and discount predictors are implemented as multilayer perceptrons (MLPs). The latent states predictor $p_\psi(\hat{z}^k_{t+1} \mid h^k_t)$ computes the parameters of a categorical distribution conditioned on the deterministic hidden state h^k_t to predict the future latent state $\hat{z}^k_{t+1}$. The reward predictor $p_\psi(\hat{r}_t \mid h'_t)$ calculates the parameters of a normal distribution conditioned on the deterministic hidden state h'_t to predict the reward $\hat{r}_t$. Both the cost predictor $p_\psi(\hat{c}_t \mid h'_t)$ and discount predictor $p_\psi(\hat{\gamma}_t \mid h'_t)$ compute the parameters of a Bernoulli distribution to predict the cost $\hat{c}_t$ and discount factor $\hat{\gamma}_t$. We visualize the trajectories generated by the model in Appendix A.

$$\begin{aligned}
&\text{Hidden state predictor:} && (h^1, \ldots, h^n)_t, h'_t = f_\psi((z^1, \ldots, z^n)_{t-l:t}, a_{t-l:t}, r_{t-l:t-1}, c_{t-l:t-1}),\\
&\text{Latent state predictor:} && \hat{z}^k_{t+1} \sim p_\psi(\hat{z}^k_{t+1} | h^k_t), \quad \text{for } k = 1, \ldots, n,\\
&\text{Reward predictor:} && \hat{r}_t \sim p_\psi(\hat{r}_t | h'_t),\\
&\text{Cost predictor:} && \hat{c}_t \sim p_\psi(\hat{c}_t | h'_t),\\
&\text{Discount predictor:} && \hat{\gamma}_t \sim p_\psi(\hat{\gamma}_t | h'_t).
\end{aligned} \tag{2}$$

Loss Functions: The goal for the slot-based autoencoder model is to minimize the reconstruction error $\mathcal{J}^t_{rec}$, while $\mathcal{J}^t_{ent}$ and $\mathcal{J}^t_{cross}$ serve as regularization terms. The entropies $\mathcal{J}^t_{ent}$ prevent the latent states $(z^1, \ldots, z^n)_t$ from collapsing into one-hot distributions. The corss-entropies $\mathcal{J}^t_{cross}$ align the encoded latent states $(z^1, \ldots, z^n)_t$ with the predicted latent states $(\hat{z}^1, \ldots, \hat{z}^n)_t$ from the dynamics model [Robine et al., 2023]. Here, since the latent states $(z^1, \ldots, z^n)_t$ encoded by the slot encoder are ordered randomly, $\mathcal{J}^t_{cross}$ is computed by rearranging them to minimize $\sum_{k=1}^n |z^k_t - \hat{z}^k_t|$.

$$\begin{aligned}
\mathcal{L}^{ae}_\phi = \mathbb{E}\left[\sum_{t=1}^T \left(\mathcal{J}^t_{rec} + \alpha_1 \mathcal{J}^t_{ent} + \alpha_2 \mathcal{J}^t_{cross}\right)\right], \quad \text{where} \quad \mathcal{J}^t_{rec} = -\ln p_\phi(o_t | (z^1, \ldots, z^n)_t),\\
\mathcal{J}^t_{ent} = -\sum_{k=1}^n H(p_\phi(z^k_t | o_t)), \quad \mathcal{J}^t_{cross} = \sum_{k=1}^n H(p_\phi(z^k_t | o_t), p_\psi(\hat{z}^k_t | h^k_{t-1})),
\end{aligned} \tag{3}$$

with hyperparameters $\alpha_1, \alpha_2 \geq 0$.

The term $\mathcal{J}^{t+1}_{cross}$, defined in Equation 3, is also used make the Transformer-based dynamics model to learn the transitions of the latent states $(z^1, \ldots, z^n)_t$. Here, $\mathcal{J}^{t+1}_{cross}$ is computed by rearranging the predicted latent states $(\hat{z}^1, \ldots, \hat{z}^n)_{t+1}$ to minimize $\sum_{k=1}^n |z^k_t - \hat{z}^k_t|$. The terms $\mathcal{J}^t_{reward}$, $\mathcal{J}^t_{cost}$, and $\mathcal{J}^t_{discount}$ are negative log-likelihoods used to optimize the reward $\hat{r}_t$, cost $\hat{c}_t$, and discount factor $\hat{\gamma}_t$, respectively. Then, the total loss functions for the Transformer-based dynamics model is givend by:

$$\begin{aligned}
\mathcal{L}^{dyn}_\psi = \mathbb{E}\left[\sum_{t=1}^T \left(\mathcal{J}^{t+1}_{cross} + \beta_1 \mathcal{J}^t_{reward} + \beta_2 \mathcal{J}^t_{cost} + \beta_3 \mathcal{J}^t_{discount}\right)\right], \quad \text{where}\\
\mathcal{J}^t_{reward} = -\ln(p_\psi(r_t | h'_t)), \mathcal{J}^t_{cost} = -\ln(p_\psi(c_t | h'_t)), \mathcal{J}^t_{discount} = -\ln(p_\psi(\gamma_t | h'_t)),
\end{aligned} \tag{4}$$

with hyperparameters $\beta_1, \beta_2, \beta_3 \geq 0$ and $\gamma_t = 0$ for episode ends and $\gamma_t = \gamma$ otherwise.

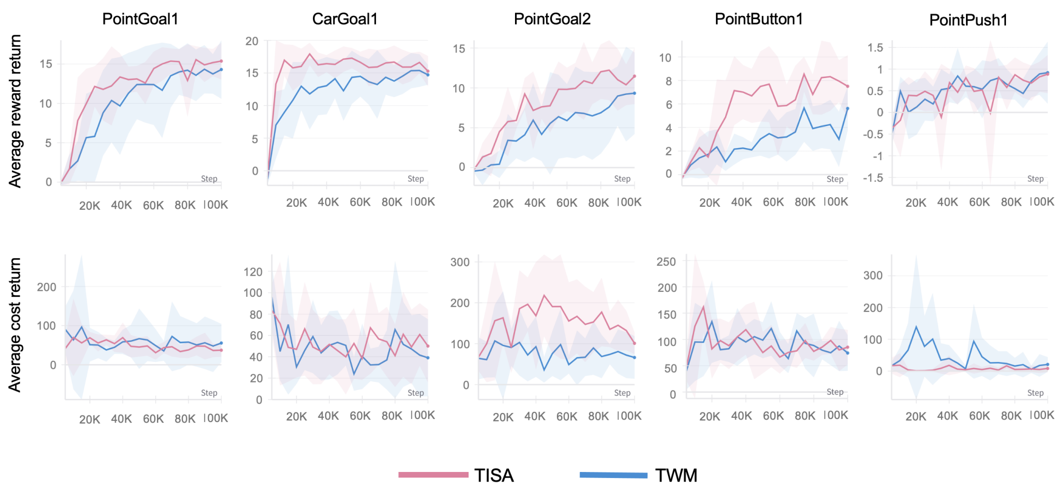


Figure 3: The TISA training results.

2.2 Transformer-based Policy and Value Functions

The policy function $\pi_\theta(a_t|\hat{z}_t^1, \ldots, \hat{z}_t^n)$, along with the two state value functions–the reward value function $v_{\xi_r}(\hat{z}_t^1, \ldots, \hat{z}_t^n)$ and the cost value function $v_{\xi_c}(\hat{z}_t^1, \ldots, \hat{z}_t^n)$–are implemented using Transformers (see Figure 1 (b)). The Transformer takes as input tokens consisting of the latent states $\hat{z}_t^1, \ldots, \hat{z}_t^n$ along with initial values, such as the action a_{init} and state values $v_{r,init}$, $v_{c,init}$, sampled from a learnable normal distribution. The Transformer outputs corresponding to these initial values are the action a_t and the state values v_r and v_c. Positional encoding is added uniformly to the latent states $\hat{z}_t^1, \ldots, \hat{z}_t^n$, ensuring that the Transformer's output does not depend on their order, making the output equivariant to the order of the latent states.

3 Experiments on Safety Gym

We have evaluated TISA on the five environments of the Safety-Gym benchmark [Achiam and Amodei, 2019] (refer to Appendix B for details). In these environments, the goal is to navigate robots to designated locations while avoiding collisions with surrounding objects. Given that Safety Gym is designed to evaluate the performance of RL safetiness, where the frequency of critical failure (the cost) is the most important metrics as well as the reward, we adopted a learning method to build a safe policy, Augmented Lagrangian method [As et al., 2022], to TISA and a comparison method (see Appendix C for details). We used performance scores used in [Achiam and Amodei, 2019] for $E = 10$ episodes of $T_{\text{ep}} = 1000$ steps. The average undiscounted reward return is defined as $\hat{J}_r(\pi) = \frac{1}{E}\sum_{i=1}^{E}\sum_{t=0}^{T_{\text{ep}}} r_t$, and the average undiscounted cost return is defined as $\hat{J}_c(\pi) = \frac{1}{E}\sum_{i=1}^{E}\sum_{t=0}^{T_{\text{ep}}} c_t$. We compared TISA with TWM [Robine et al., 2023], a Transformer-based world model method without object-centric representations. We limited interactions with the environments within 100K and obtained the average scores over five runs. Experimental details are provided in Appendix D.

4 Results

We summarized the time courses of the scores in Figure 3. The solid lines denote the average over five runs, while the shaded areas denote the standard deviations. In three out of the five environments–PointGoal1, CarGoal1 and PointButton1–TISA has achieved the better rewards with the costs at the same levels as TWM. The success of TISA demonstrates that the Transformer-based world model with object-centric representations effectively captures the dynamics of individual objects, leading to more efficient learning. Additionally, the Transformer-based policy and value functions promote the learning of more effective policies by taking into account the unique dynamics of each object. This is particularly evident in the PointButton1 environment, which features the greatest variety of objects among the five environments, including dynamic obstacles and the most diverse object shapes. Here, TISA shows a clear performance advantage over TWM, highlighting its capability to learn effectively even in complex environments. Also, the limitations of TISA are discussed in Appendix E.

References

Joshua Achiam and Dario Amodei. Benchmarking safe exploration in deep reinforcement learning. 2019.

Yarden As, Ilnura Usmanova, Sebastian Curi, and Andreas Krause. Constrained policy optimization via bayesian world models. In *International Conference on Learning Representations*, 2022.

David Ha and Jürgen Schmidhuber. Recurrent world models facilitate policy evolution. In *Advances in Neural Information Processing Systems*, 2018.

Danijar Hafner, Timothy Lillicrap, Jimmy Ba, and Mohammad Norouzi. Dream to control: Learning behaviors by latent imagination. In *International Conference on Learning Representations*, 2020.

Danijar Hafner, Timothy P Lillicrap, Mohammad Norouzi, and Jimmy Ba. Mastering atari with discrete world models. In *International Conference on Learning Representations*, 2021.

Dan Haramati, Tal Daniel, and Aviv Tamar. Entity-centric reinforcement learning for object manipulation from pixels. In *International Conference on Learning Representations*, 2024.

Zhixuan Lin, Yi-Fu Wu, Skand Peri, Bofeng Fu, Jindong Jiang, and Sungjin Ahn. Improving generative imagination in object-centric world models. In *International Conference on Machine Learning*, 2020.

Francesco Locatello, Dirk Weissenborn, Thomas Unterthiner, Aravindh Mahendran, Georg Heigold, Jakob Uszkoreit, Alexey Dosovitskiy, and Thomas Kipf. Object-centric learning with slot attention. In *Advances in Neural Information Processing Systems*, 2020.

Michael McCloskey, Allyson Washburn, and Linda Felch. Intuitive physics: The straight-down belief and its origin. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 9:636–649, 1983.

Vincent Micheli, Eloi Alonso, and François Fleuret. Transformers are sample-efficient world models. In *International Conference on Learning Representations*, 2023.

Volodymyr Mnih, Adrià Puigdomènech Badia, Mehdi Mirza, Alex Graves, Timothy P. Lillicrap, Tim Harley, David Silver, and Koray Kavukcuoglu. Asynchronous methods for deep reinforcement learning. In *International Conference on Machine Learning*, 2016.

Jan Robine, Marc Höftmann, Tobias Uelwer, and Stefan Harmeling. Transformer-based world models are happy with 100k interactions. In *International Conference on Learning Representations*, 2023.

John Schulman, Philipp Moritz, Sergey Levine, Michael Jordan, and Pieter Abbeel. High-dimensional continuous control using generalized advantage estimation. In *International Conference on Learning Representations*, 2016.

Elizabeth S. Spelke and Katherine D. Kinzler. Core knowledge. *Developmental Science*, 10(1):89–96, 2007. doi: 10.1111/j.1467-7687.2007.00569.x.

Nicholas Watters, Loïc Matthey, Christopher P. Burgess, and Alexander Lerchner. Spatial broadcast decoder: A simple architecture for learning disentangled representations in vaes. *ArXiv*, abs/1901.07017, 2019.

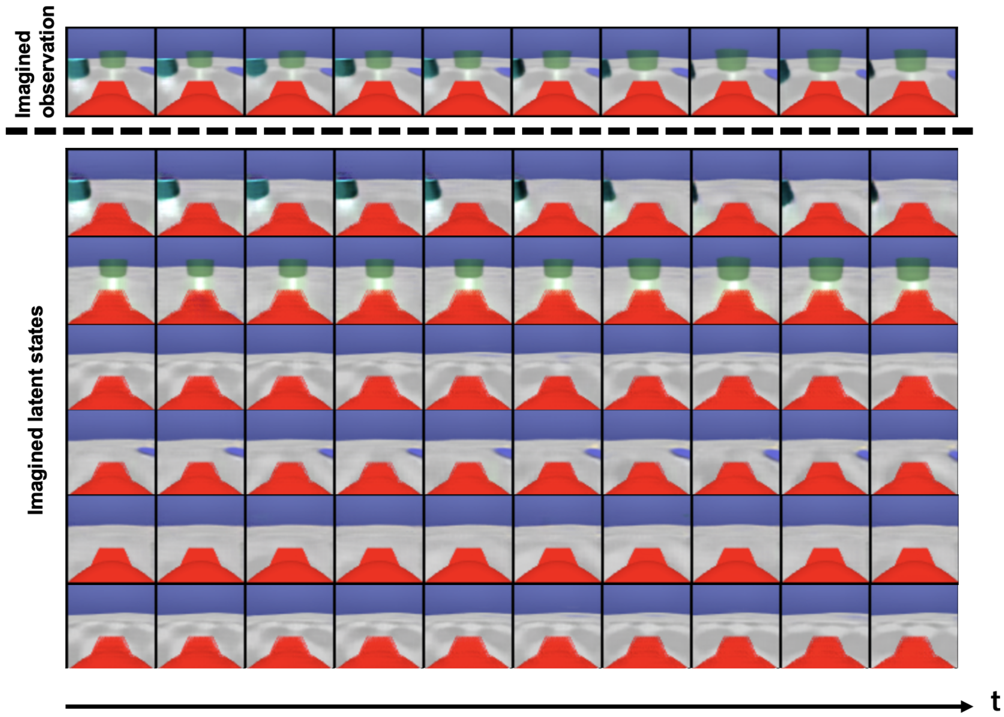
Ziyi Wu, Nikita Dvornik, Klaus Greff, Thomas Kipf, and Animesh Garg. Slotformer: Unsupervised visual dynamics simulation with object-centric models. In *International Conference on Learning Representations*, 2023.

Jaesik Yoon, Yi-Fu Wu, Heechul Bae, and Sungjin Ahn. An investigation into pre-training object-centric representations for reinforcement learning. In *International Conference on Machine Learning*, 2023.

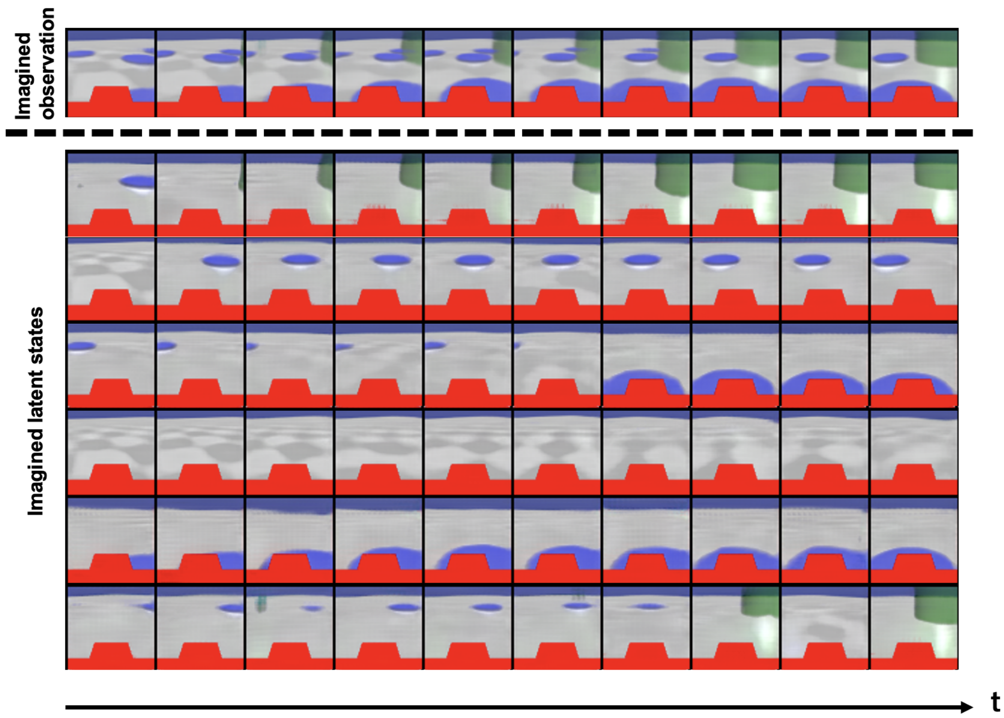
Weipu Zhang, Gang Wang, Jian Sun, Yetian Yuan, and Gao Huang. STORM: Efficient stochastic transformer based world models for reinforcement learning. In *Advances in Neural Information Processing Systems*, 2023.

Daniel Zoran, Rishabh Kabra, Alexander Lerchner, and Danilo Jimenez Rezende. Parts: Unsupervised segmentation with slots, attention and independence maximization. *2021 IEEE/CVF International Conference on Computer Vision*, pages 10419–10427, 2021.

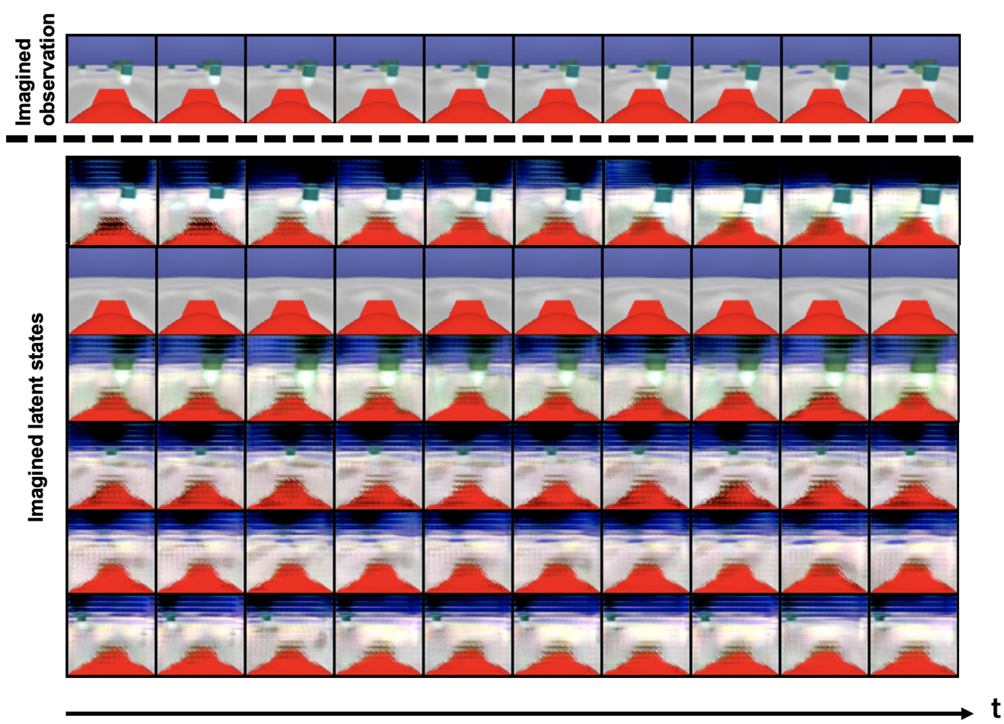
A Trajectories generated by our world model



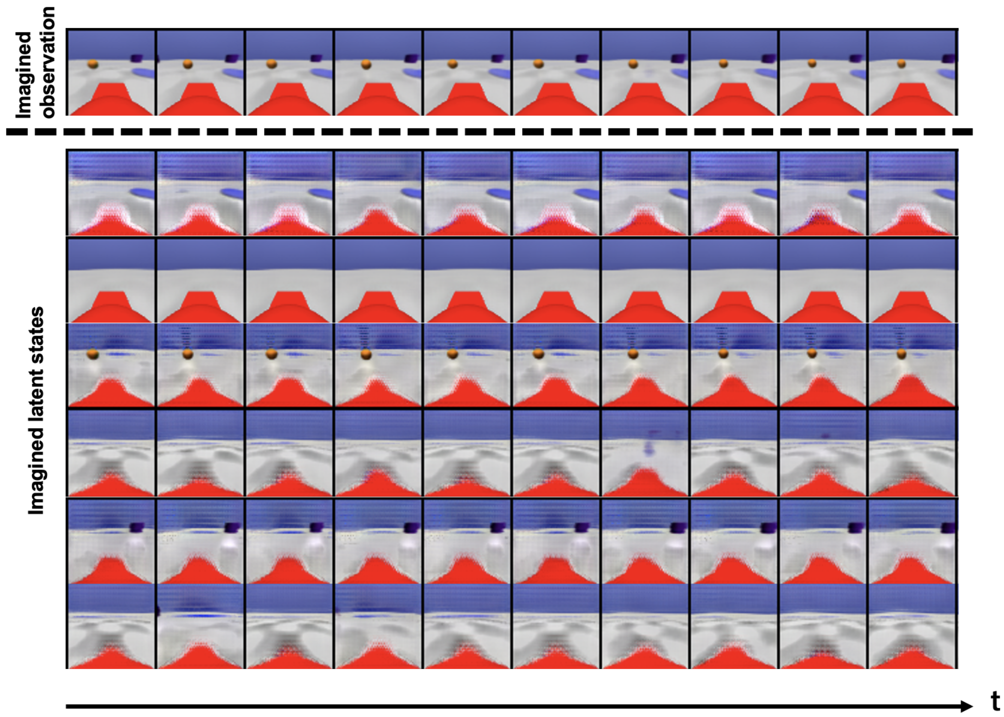
(a) PointGoal1.



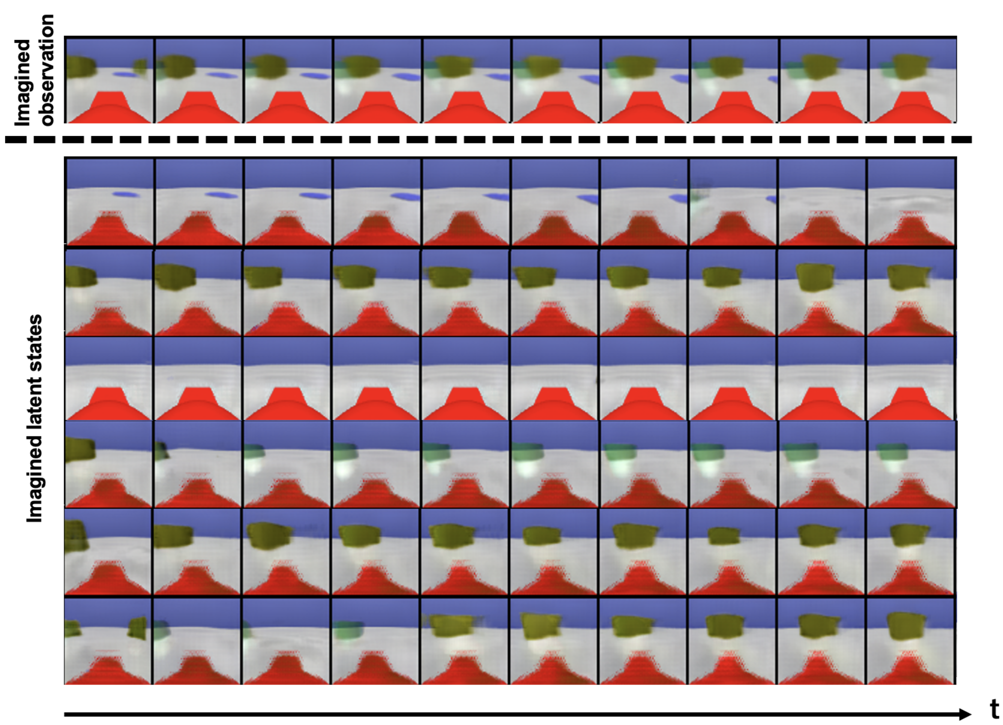
(b) CarGoal1.



(c) PointGoal2.

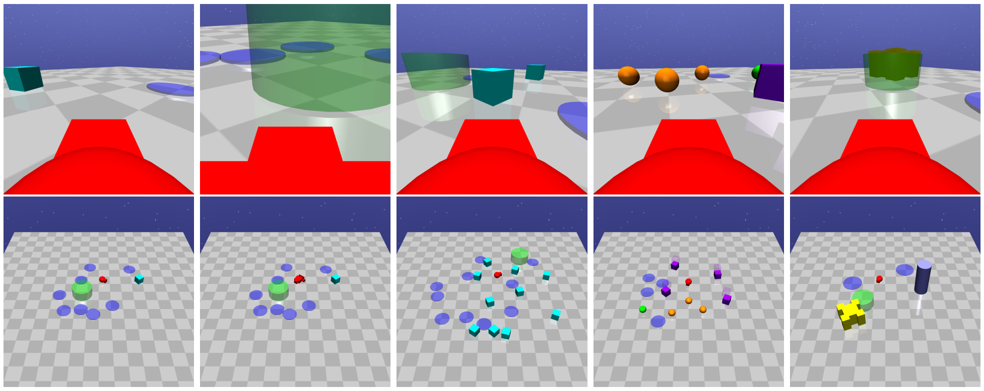


(d) PointButton1.



(e) PointPush1.

B Safety Gym Benchmark Tasks



PointGoal1 CarGoal1 PointGoal2 PointButton1 PointPush1

Figure 4: 5 tasks in Safety Gym. In our experiments, we utilize first-person-view images with a resolution of 64×64 pixels as observations. The green objects indicate the goals the robot needs to reach. In the PointPush1 task, the yellow box must be pushed into the goal area, while interacting with any other object type is regarded as unsafe behavior.

C Policy Objective Function

The objective of the safe reinforcement learning problem is to maximize the accumulated reward under a given safety threshold b:

$$\max_{\pi} \mathbb{E}_{\pi}\left[\sum_{t=0}^{\infty} \gamma^t r_t\right] \quad \text{s.t.} \quad \mathbb{E}_{\pi}\left[\sum_{t=0}^{\infty} \gamma^t c_t\right] \leq b. \tag{5}$$

To solve this problem, following the Augmented Lagrangian method from [As et al., 2022], we define the objective function of the policy as follows:

$$\mathcal{J}_{\theta} = \mathcal{J}_{\theta}^{r} - \Psi(\mathcal{J}_{\theta}^{c}, \lambda_k, \mu_k), \tag{6}$$

where

$$\begin{aligned} \mathcal{J}_{\theta}^{r} &= \mathbb{E}_{\pi_{\theta}}\left[\ln \pi_{\theta}(a_t|\hat{z}_t^1, \ldots, \hat{z}_t^n) A^r(\hat{z}_t^1, \ldots, \hat{z}_t^n, a_t)\right], \\ \mathcal{J}_{\theta}^{c} &= \mathbb{E}_{\pi_{\theta}}\left[\ln \pi_{\theta}(a_t|\hat{z}_t^1, \ldots, \hat{z}_t^n) A^c(\hat{z}_t^1, \ldots, \hat{z}_t^n, a_t)\right]. \end{aligned} \tag{7}$$

Here, $A^r(\hat{z}_t^1, \ldots, \hat{z}_t^n, a_t)$ is the advantage function [Mnih et al., 2016] for the reward, and $A^c(\hat{z}_t^1, \ldots, \hat{z}_t^n, a_t)$ is the advantage function for the cost, computed similarly to the reward's advantage function. Both advantage functions are estimated using the generalized advantage estimator (GAE) [Schulman et al., 2016]. Additionally, $\mu_k = \max(\mu_{k-1}(\nu + 1.0), 1.0)$ represents a monotonically non-decreasing term corresponding to the gradient step k, where $\nu > 0$. The penalty term and Lagrange multiplier in the loss function are updated as follows:

$$\Psi(\mathcal{J}_{\theta}^{c}, \lambda_k, \mu_k), \lambda_p^{k+1} = \begin{cases} \lambda_k(\mathcal{J}_{\theta}^{c} - b) + \mu_k(\mathcal{J}_{\theta}^{c} - b)^2, \lambda_k + \mu_k(\mathcal{J}_{\theta}^{c} - b) & \text{if } \lambda_k + \mu_k(\mathcal{J}_{\theta}^{c} - b) \geq 0 \\ -\frac{(\lambda_k)^2}{2\mu_k}, 0 & \text{otherwise.} \end{cases} \tag{8}$$

D Experimental Details

TISA takes approximately 4 days to train with 2 NVIDIA V100 GPUs. The hyperparameters used in the experiments are listed in Table 2.

Table 2: Hyperparameters

Name	Symbol	Value
Number of Slots	n	6
Slot Size		128
Number of Slot Attention iterations		2
World Model Batch Size		64
History Length	l	16
Imagination Batch Size		400
Imagination Horizon		15
Discount Factor	γ	0.99
GAE Parameter		0.95
Penalty Term	ν	1e-5
Initial Penalty Multiplier	μ_0	0.95
Initial Lagrangian Multiplier	λ_0	2e-4
Safe Threshold	b	-0.1
Coefficient of $\mathcal{J}_{ent}$	α_1	5.0
Coefficient $\mathcal{J}_{cross}$	α_2	0.01
Coefficient of $\mathcal{J}_{reward}$	β_1	10.0
Coefficient of $\mathcal{J}_{cost}$	β_2	50.0
Coefficient of $\mathcal{J}_{discount}$	β_3	50.0
Environment Steps	-	100K
Action Repeat	-	2
Observation Learing Rate	-	0.0001
Dynamics Learning Rate	-	0.0001
Actor Learning Rate	-	0.0001
Critic Learning Rate	-	0.00001
Safety Critic Leaning Rate	-	0.00001

E Limitations

One issue of TISA is computational efficiency. Compared to TWM, TISA requires twice the computational resources and eight times longer run time. However, as shown in Figure 3, it demonstrates better sample efficiency than TWM.

Another issue is that the autoencoder used in TISA employs the simplest decoder among those used in slot-based object-centric representation learning methods, namely the Mixture Decoder [Locatello et al., 2020], which does not work well with naturalistic images. However, this problem could be addressed by using a more advanced decoder that is better suited for handling naturalistic images.

NeurIPS Paper Checklist

The checklist is designed to encourage best practices for responsible machine learning research, addressing issues of reproducibility, transparency, research ethics, and societal impact. Do not remove the checklist: **The papers not including the checklist will be desk rejected.** The checklist should follow the references and follow the (optional) supplemental material. The checklist does NOT count towards the page limit.

Please read the checklist guidelines carefully for information on how to answer these questions. For each question in the checklist:

- You should answer [Yes] , [No] , or [NA] .
- [NA] means either that the question is Not Applicable for that particular paper or the relevant information is Not Available.
- Please provide a short (1–2 sentence) justification right after your answer (even for NA).

The checklist answers are an integral part of your paper submission. They are visible to the reviewers, area chairs, senior area chairs, and ethics reviewers. You will be asked to also include it (after eventual revisions) with the final version of your paper, and its final version will be published with the paper.

The reviewers of your paper will be asked to use the checklist as one of the factors in their evaluation. While "[Yes] " is generally preferable to "[No] ", it is perfectly acceptable to answer "[No] " provided a proper justification is given (e.g., "error bars are not reported because it would be too computationally expensive" or "we were unable to find the license for the dataset we used"). In general, answering "[No] " or "[NA] " is not grounds for rejection. While the questions are phrased in a binary way, we acknowledge that the true answer is often more nuanced, so please just use your best judgment and write a justification to elaborate. All supporting evidence can appear either in the main paper or the supplemental material, provided in appendix. If you answer [Yes] to a question, in the justification please point to the section(s) where related material for the question can be found.

IMPORTANT, please:

- **Delete this instruction block, but keep the section heading "NeurIPS paper checklist"**,
- **Keep the checklist subsection headings, questions/answers and guidelines below.**
- **Do not modify the questions and only use the provided macros for your answers**.

1. **Claims**

   Question: Do the main claims made in the abstract and introduction accurately reflect the paper's contributions and scope?

   Answer: [Yes]

   Justification: It is shown in Section 4, Results.

   Guidelines:

   - The answer NA means that the abstract and introduction do not include the claims made in the paper.
   - The abstract and/or introduction should clearly state the claims made, including the contributions made in the paper and important assumptions and limitations. A No or NA answer to this question will not be perceived well by the reviewers.
   - The claims made should match theoretical and experimental results, and reflect how much the results can be expected to generalize to other settings.
   - It is fine to include aspirational goals as motivation as long as it is clear that these goals are not attained by the paper.

2. **Limitations**

   Question: Does the paper discuss the limitations of the work performed by the authors?

   Answer: [Yes]

   Justification: It is detailed in the Limitations section of the Appendix.

   Guidelines:

   - The answer NA means that the paper has no limitation while the answer No means that the paper has limitations, but those are not discussed in the paper.
   - The authors are encouraged to create a separate "Limitations" section in their paper.

- The paper should point out any strong assumptions and how robust the results are to violations of these assumptions (e.g., independence assumptions, noiseless settings, model well-specification, asymptotic approximations only holding locally). The authors should reflect on how these assumptions might be violated in practice and what the implications would be.
- The authors should reflect on the scope of the claims made, e.g., if the approach was only tested on a few datasets or with a few runs. In general, empirical results often depend on implicit assumptions, which should be articulated.
- The authors should reflect on the factors that influence the performance of the approach. For example, a facial recognition algorithm may perform poorly when image resolution is low or images are taken in low lighting. Or a speech-to-text system might not be used reliably to provide closed captions for online lectures because it fails to handle technical jargon.
- The authors should discuss the computational efficiency of the proposed algorithms and how they scale with dataset size.
- If applicable, the authors should discuss possible limitations of their approach to address problems of privacy and fairness.
- While the authors might fear that complete honesty about limitations might be used by reviewers as grounds for rejection, a worse outcome might be that reviewers discover limitations that aren't acknowledged in the paper. The authors should use their best judgment and recognize that individual actions in favor of transparency play an important role in developing norms that preserve the integrity of the community. Reviewers will be specifically instructed to not penalize honesty concerning limitations.

3. **Theory Assumptions and Proofs**

   Question: For each theoretical result, does the paper provide the full set of assumptions and a complete (and correct) proof?

   Answer: [NA]

   Justification: The paper does not include theoretical results.

   Guidelines:

   - The answer NA means that the paper does not include theoretical results.
   - All the theorems, formulas, and proofs in the paper should be numbered and cross-referenced.
   - All assumptions should be clearly stated or referenced in the statement of any theorems.
   - The proofs can either appear in the main paper or the supplemental material, but if they appear in the supplemental material, the authors are encouraged to provide a short proof sketch to provide intuition.
   - Inversely, any informal proof provided in the core of the paper should be complemented by formal proofs provided in appendix or supplemental material.
   - Theorems and Lemmas that the proof relies upon should be properly referenced.

4. **Experimental Result Reproducibility**

   Question: Does the paper fully disclose all the information needed to reproduce the main experimental results of the paper to the extent that it affects the main claims and/or conclusions of the paper (regardless of whether the code and data are provided or not)?

   Answer: [Yes]

   Justification: It is detailed in the Experiments section of Section 4 and the Experimental Details section of the Appendix.

   Guidelines:

   - The answer NA means that the paper does not include experiments.
   - If the paper includes experiments, a No answer to this question will not be perceived well by the reviewers: Making the paper reproducible is important, regardless of whether the code and data are provided or not.
   - If the contribution is a dataset and/or model, the authors should describe the steps taken to make their results reproducible or verifiable.
   - Depending on the contribution, reproducibility can be accomplished in various ways. For example, if the contribution is a novel architecture, describing the architecture fully might suffice, or if the contribution is a specific model and empirical evaluation, it may be necessary to either make it possible for others to replicate the model with the same dataset, or provide access to the model. In general. releasing code and data is often one good way to accomplish this, but reproducibility can also be provided via detailed instructions for how to replicate the results, access to a hosted model (e.g., in the case of a large language model), releasing of a model checkpoint, or other means that are appropriate to the research performed.

- While NeurIPS does not require releasing code, the conference does require all submissions to provide some reasonable avenue for reproducibility, which may depend on the nature of the contribution. For example
  (a) If the contribution is primarily a new algorithm, the paper should make it clear how to reproduce that algorithm.
  (b) If the contribution is primarily a new model architecture, the paper should describe the architecture clearly and fully.
  (c) If the contribution is a new model (e.g., a large language model), then there should either be a way to access this model for reproducing the results or a way to reproduce the model (e.g., with an open-source dataset or instructions for how to construct the dataset).
  (d) We recognize that reproducibility may be tricky in some cases, in which case authors are welcome to describe the particular way they provide for reproducibility. In the case of closed-source models, it may be that access to the model is limited in some way (e.g., to registered users), but it should be possible for other researchers to have some path to reproducing or verifying the results.

5. **Open access to data and code**

   Question: Does the paper provide open access to the data and code, with sufficient instructions to faithfully reproduce the main experimental results, as described in supplemental material?

   Answer: [No]

   Justification: Unfortunately, we were unable to prepare the data and code for open access in time for this submission. We will make the necessary resources available and provide detailed instructions to reproduce the main experimental results as soon as possible.

   Guidelines:

   - The answer NA means that paper does not include experiments requiring code.
   - Please see the NeurIPS code and data submission guidelines (`https://nips.cc/public/guides/CodeSubmissionPolicy`) for more details.
   - While we encourage the release of code and data, we understand that this might not be possible, so "No" is an acceptable answer. Papers cannot be rejected simply for not including code, unless this is central to the contribution (e.g., for a new open-source benchmark).
   - The instructions should contain the exact command and environment needed to run to reproduce the results. See the NeurIPS code and data submission guidelines (`https://nips.cc/public/guides/CodeSubmissionPolicy`) for more details.
   - The authors should provide instructions on data access and preparation, including how to access the raw data, preprocessed data, intermediate data, and generated data, etc.
   - The authors should provide scripts to reproduce all experimental results for the new proposed method and baselines. If only a subset of experiments are reproducible, they should state which ones are omitted from the script and why.
   - At submission time, to preserve anonymity, the authors should release anonymized versions (if applicable).
   - Providing as much information as possible in supplemental material (appended to the paper) is recommended, but including URLs to data and code is permitted.

6. **Experimental Setting/Details**

   Question: Does the paper specify all the training and test details (e.g., data splits, hyperparameters, how they were chosen, type of optimizer, etc.) necessary to understand the results?

   Answer: [Yes]

   Justification: It is detailed in the Experimental Details section of the Appendix.

   Guidelines:

   - The answer NA means that the paper does not include experiments.
   - The experimental setting should be presented in the core of the paper to a level of detail that is necessary to appreciate the results and make sense of them.
   - The full details can be provided either with the code, in appendix, or as supplemental material.

7. **Experiment Statistical Significance**

   Question: Does the paper report error bars suitably and correctly defined or other appropriate information about the statistical significance of the experiments?

   Answer: [Yes]

   Justification: It is used in the Results section of Section 4, and its explanation is clearly provided.

Guidelines:

- The answer NA means that the paper does not include experiments.
- The authors should answer "Yes" if the results are accompanied by error bars, confidence intervals, or statistical significance tests, at least for the experiments that support the main claims of the paper.
- The factors of variability that the error bars are capturing should be clearly stated (for example, train/test split, initialization, random drawing of some parameter, or overall run with given experimental conditions).
- The method for calculating the error bars should be explained (closed form formula, call to a library function, bootstrap, etc.)
- The assumptions made should be given (e.g., Normally distributed errors).
- It should be clear whether the error bar is the standard deviation or the standard error of the mean.
- It is OK to report 1-sigma error bars, but one should state it. The authors should preferably report a 2-sigma error bar than state that they have a 96% CI, if the hypothesis of Normality of errors is not verified.
- For asymmetric distributions, the authors should be careful not to show in tables or figures symmetric error bars that would yield results that are out of range (e.g. negative error rates).
- If error bars are reported in tables or plots, The authors should explain in the text how they were calculated and reference the corresponding figures or tables in the text.

8. **Experiments Compute Resources**

   Question: For each experiment, does the paper provide sufficient information on the computer resources (type of compute workers, memory, time of execution) needed to reproduce the experiments?

   Answer: [Yes]

   Justification: This is detailed in the Experimental Details section of the Appendix.

   Guidelines:

   - The answer NA means that the paper does not include experiments.
   - The paper should indicate the type of compute workers CPU or GPU, internal cluster, or cloud provider, including relevant memory and storage.
   - The paper should provide the amount of compute required for each of the individual experimental runs as well as estimate the total compute.
   - The paper should disclose whether the full research project required more compute than the experiments reported in the paper (e.g., preliminary or failed experiments that didn't make it into the paper).

9. **Code Of Ethics**

   Question: Does the research conducted in the paper conform, in every respect, with the NeurIPS Code of Ethics `https://neurips.cc/public/EthicsGuidelines`?

   Answer: [Yes]

   Justification: This research is fully compliant with the NeurIPS Code of Ethics.

   Guidelines:

   - The answer NA means that the authors have not reviewed the NeurIPS Code of Ethics.
   - If the authors answer No, they should explain the special circumstances that require a deviation from the Code of Ethics.
   - The authors should make sure to preserve anonymity (e.g., if there is a special consideration due to laws or regulations in their jurisdiction).

10. **Broader Impacts**

    Question: Does the paper discuss both potential positive societal impacts and negative societal impacts of the work performed?

    Answer: [No]

    Justification: The paper does not discuss both potential positive and negative societal impacts of the work performed.

    Guidelines:

    - The answer NA means that there is no societal impact of the work performed.
    - If the authors answer NA or No, they should explain why their work has no societal impact or why the paper does not address societal impact.

- Examples of negative societal impacts include potential malicious or unintended uses (e.g., disinformation, generating fake profiles, surveillance), fairness considerations (e.g., deployment of technologies that could make decisions that unfairly impact specific groups), privacy considerations, and security considerations.
- The conference expects that many papers will be foundational research and not tied to particular applications, let alone deployments. However, if there is a direct path to any negative applications, the authors should point it out. For example, it is legitimate to point out that an improvement in the quality of generative models could be used to generate deepfakes for disinformation. On the other hand, it is not needed to point out that a generic algorithm for optimizing neural networks could enable people to train models that generate Deepfakes faster.
- The authors should consider possible harms that could arise when the technology is being used as intended and functioning correctly, harms that could arise when the technology is being used as intended but gives incorrect results, and harms following from (intentional or unintentional) misuse of the technology.
- If there are negative societal impacts, the authors could also discuss possible mitigation strategies (e.g., gated release of models, providing defenses in addition to attacks, mechanisms for monitoring misuse, mechanisms to monitor how a system learns from feedback over time, improving the efficiency and accessibility of ML).

11. **Safeguards**

    Question: Does the paper describe safeguards that have been put in place for responsible release of data or models that have a high risk for misuse (e.g., pretrained language models, image generators, or scraped datasets)?

    Answer: [NA]

    Justification: The paper poses no such risks.

    Guidelines:

    - The answer NA means that the paper poses no such risks.
    - Released models that have a high risk for misuse or dual-use should be released with necessary safeguards to allow for controlled use of the model, for example by requiring that users adhere to usage guidelines or restrictions to access the model or implementing safety filters.
    - Datasets that have been scraped from the Internet could pose safety risks. The authors should describe how they avoided releasing unsafe images.
    - We recognize that providing effective safeguards is challenging, and many papers do not require this, but we encourage authors to take this into account and make a best faith effort.

12. **Licenses for existing assets**

    Question: Are the creators or original owners of assets (e.g., code, data, models), used in the paper, properly credited and are the license and terms of use explicitly mentioned and properly respected?

    Answer: [Yes]

    Justification: All methods used in this paper are appropriately cited within the paper.

    Guidelines:

    - The answer NA means that the paper does not use existing assets.
    - The authors should cite the original paper that produced the code package or dataset.
    - The authors should state which version of the asset is used and, if possible, include a URL.
    - The name of the license (e.g., CC-BY 4.0) should be included for each asset.
    - For scraped data from a particular source (e.g., website), the copyright and terms of service of that source should be provided.
    - If assets are released, the license, copyright information, and terms of use in the package should be provided. For popular datasets, `paperswithcode.com/datasets` has curated licenses for some datasets. Their licensing guide can help determine the license of a dataset.
    - For existing datasets that are re-packaged, both the original license and the license of the derived asset (if it has changed) should be provided.
    - If this information is not available online, the authors are encouraged to reach out to the asset's creators.

13. **New Assets**

    Question: Are new assets introduced in the paper well documented and is the documentation provided alongside the assets?

    Answer: [Yes]

Justification:In Section 3, we propose a new model and provide a detailed explanation of it.

Guidelines:

- The answer NA means that the paper does not release new assets.
- Researchers should communicate the details of the dataset/code/model as part of their submissions via structured templates. This includes details about training, license, limitations, etc.
- The paper should discuss whether and how consent was obtained from people whose asset is used.
- At submission time, remember to anonymize your assets (if applicable). You can either create an anonymized URL or include an anonymized zip file.

14. **Crowdsourcing and Research with Human Subjects**

    Question: For crowdsourcing experiments and research with human subjects, does the paper include the full text of instructions given to participants and screenshots, if applicable, as well as details about compensation (if any)?

    Answer: [NA]

    Justification: The paper does not involve crowdsourcing nor research with human subjects.

    Guidelines:

    - The answer NA means that the paper does not involve crowdsourcing nor research with human subjects.
    - Including this information in the supplemental material is fine, but if the main contribution of the paper involves human subjects, then as much detail as possible should be included in the main paper.
    - According to the NeurIPS Code of Ethics, workers involved in data collection, curation, or other labor should be paid at least the minimum wage in the country of the data collector.

15. **Institutional Review Board (IRB) Approvals or Equivalent for Research with Human Subjects**

    Question: Does the paper describe potential risks incurred by study participants, whether such risks were disclosed to the subjects, and whether Institutional Review Board (IRB) approvals (or an equivalent approval/review based on the requirements of your country or institution) were obtained?

    Answer: [NA]

    Justification: The paper does not involve crowdsourcing nor research with human subjects.

    Guidelines:

    - The answer NA means that the paper does not involve crowdsourcing nor research with human subjects.
    - Depending on the country in which research is conducted, IRB approval (or equivalent) may be required for any human subjects research. If you obtained IRB approval, you should clearly state this in the paper.
    - We recognize that the procedures for this may vary significantly between institutions and locations, and we expect authors to adhere to the NeurIPS Code of Ethics and the guidelines for their institution.
    - For initial submissions, do not include any information that would break anonymity (if applicable), such as the institution conducting the review.