

Planificación por descenso de gradiente a partir de observaciones visuales

Vélez García, C.^{a,*}, Cazorla, Miguel.^b, Pomares, Jorge.^b

^aINESCOP, C/ Alemania, 102, Elda, 03600 Alicante, España.

^bUniversity Institute for Computing Research, Universidad de Alicante, San Vicente del Raspeig, 03690, Alicante, España.

Resumen

La planificación de trayectorias condicionada a objetivos en robótica se beneficia de operar en espacios latentes compactos y semánticos. Sin embargo, gran parte de la literatura asume entradas en espacio de estados, mientras que muchos escenarios reales dependen de observaciones visuales de alta dimensión. En este trabajo preliminar proponemos una extensión de la planificación por descenso (Planning-as-Descent) al dominio visual, evitando planificar directamente en espacio de píxeles. La idea central es aprender una representación visual semántica mediante auto-supervisión (tokenización convolucional dispersa y modelado enmascarado predictivo en espacio latente) y reutilizar la formulación energética y el procedimiento de refinamiento por gradiente del planificador en el espacio latente resultante. Presentamos la formulación, las hipótesis y un protocolo experimental reproducible para evaluar (i) estabilidad del descenso energético, (ii) eficiencia de planificación y (iii) robustez al replanteamiento, comparando con variantes basadas en estados y en píxeles. Este manuscrito define la integración y el diseño experimental como base para validaciones en curso.

Palabras clave: Planificación robótica, Control condicionado a objetivos, Auto-supervisión, Representación visual, Modelado enmascarado.

Planning as Descent from Visual Observations

Abstract

Goal-conditioned robot planning benefits from compact and semantic latent spaces, yet many planners assume access to low-dimensional state inputs, while real-world deployments often rely on high-dimensional visual observations. This work-in-progress proposes extending Planning-as-Descent to the visual domain without planning in pixel space. We learn semantic visual representations via self-supervision (a sparse convolutional tokenizer and masked predictive learning in latent space) and then reuse the original energy-based formulation and gradient-based plan refinement in the resulting latent space. We present the method, key design choices, and a reproducible experimental protocol to evaluate (i) energy-descent stability, (ii) planning efficiency, and (iii) replanning robustness, including comparisons against state-based and naïve pixel-based variants. No empirical results are reported due to time constraints; instead, we provide testable hypotheses and a concrete roadmap for ongoing validation.

Keywords: Robot planning, Goal-conditioned control, Self-supervised learning, Visual representation, Masked modeling.

1. Introducción

La planificación robótica condicionada a objetivos puede formularse como el problema de sintetizar una secuencia de acciones que lleve un sistema dinámico desde una configuración actual hasta un objetivo deseado, respetando restricciones implícitas de dinámica y factibilidad. Desde una perspectiva de control, este problema puede interpretarse como la búsqueda de una trayectoria en el espacio de estados que minimice un funcional asociado a coste o energía, sujeto a la dinámica del

sistema.

En los últimos años, los enfoques energéticos en espacio latente han propuesto una reinterpretación de este problema: en lugar de optimizar directamente sobre acciones o estados físicos, se aprende una representación compacta en la que la compatibilidad entre trayectorias, contexto pasado y objetivos puede modelarse mediante una función energética diferenciable (García et al., 2025b). Bajo esta formulación, la planificación se convierte en un proceso de inferencia por descenso de gradien-

*Autor para correspondencia: cvg25@alu.ua.es

te sobre trayectorias latentes. Este paradigma presenta varias ventajas teóricas y prácticas: (i) permite evaluar plausibilidad global de una trayectoria mediante un único escalar energético, (ii) unifica entrenamiento e inferencia bajo el mismo principio, y (iii) facilita el replanteamiento continuo al reinterpretar la planificación como un problema de optimización iterativa.

No obstante, la mayoría de estos enfoques asumen acceso a un espacio de estados explícito, típicamente compuesto por variables físicas como posiciones y velocidades. Esta hipótesis resulta restrictiva en aplicaciones reales. En manipulación robótica y entornos no estructurados, la percepción visual constituye con frecuencia la modalidad primaria, y el sistema opera a partir de observaciones de alta dimensión en espacio de píxeles. En este contexto, planificar directamente sobre imágenes presenta desafíos fundamentales: alta dimensionalidad, sensibilidad a variaciones de apariencia no causales (iluminación, textura, fondo) y paisajes de optimización potencialmente irregulares.

Desde una perspectiva más general, esta limitación refleja una brecha entre representación perceptiva y planificación: mientras que la teoría de control clásico opera en espacios de estados bien definidos, los sistemas perceptivos modernos producen observaciones crudas cuya relación con las variables dinámicas subyacentes es indirecta. Por tanto, surge la necesidad de un espacio intermedio que preserve la estructura semántica relevante para la tarea y, al mismo tiempo, sea adecuado para procesos de optimización energética.

En este trabajo preliminar proponemos extender el paradigma *Planning-as-Descent* al dominio visual mediante el aprendizaje previo de un espacio latente visual semántico. La hipótesis central es que, si el embedding visual captura factores causales y regularidades estructurales del entorno, entonces el descenso energético en dicho espacio puede aproximar el comportamiento observado cuando se dispone de estados explícitos, evitando al mismo tiempo las limitaciones de la optimización directa en píxel.

Las contribuciones de este manuscrito son tres: (i) formular de manera explícita la extensión visual del modelo energético de planificación, (ii) establecer un marco experimental comparativo entre planificación en estado, en píxel y en latente visual, y (iii) definir hipótesis verificables que guíen la validación empírica en curso. Este trabajo se posiciona como un paso intermedio hacia la integración coherente entre representación auto-supervisada y planificación energética en robótica basada en percepción.

2. Trabajo relacionado

2.1. Planificación condicionada a objetivos en espacio latente

La planificación condicionada a objetivos puede abordarse desde distintas perspectivas: optimización directa en espacio de acciones (Henaff et al., 2019; Chen et al., 2021), modelado autoregresivo de trayectorias (Carroll et al., 2022; Wu et al., 2023), modelos generativos basados en difusión (Janner et al., 2022; Chi et al., 2025) y enfoques energéticos (García et al., 2025b; Florence et al., 2022). En los últimos años, la representación en espacio latente ha emergido como un mecanis-

mo clave para reducir dimensionalidad y capturar regularidades dinámicas relevantes (LeCun et al., 2022).

Los planificadores en espacio latente buscan representar tanto la dinámica como los objetivos en una variable compacta que preserve la estructura del problema. Bajo esta formulación, la planificación se interpreta como inferencia en un modelo implícito de trayectorias. En los enfoques basados en energía, se aprende una función E_θ tal que trayectorias consistentes con el contexto y el objetivo reciben baja energía, mientras que trayectorias inconsistentes reciben alta energía (García et al., 2025b; Florence et al., 2022).

La inferencia se realiza mediante descenso de gradiente sobre la trayectoria candidata, aproximando modos de alta probabilidad. Este mecanismo presenta ventajas frente a enfoques autoregresivos puros, ya que permite refinamiento paralelo de la trayectoria completa y minimiza la acumulación de error paso a paso. Asimismo, a diferencia de modelos de difusión, el descenso energético puede implementarse con un número reducido de iteraciones, facilitando replanteamiento en tiempo real.

No obstante, la mayoría de estos trabajos asumen que el espacio latente se deriva de estados físicos explícitos o de representaciones ya alineadas con variables dinámicas. La extensión natural hacia observaciones visuales introduce el desafío de aprender un espacio latente que preserve la información causal necesaria para planificación sin heredar la complejidad del espacio de píxeles.

2.2. Representación visual auto-supervisada

El aprendizaje auto-supervisado en visión ha demostrado que es posible extraer representaciones transferibles sin anotaciones manuales. Entre los enfoques más influyentes se encuentran los modelos de reconstrucción enmascarada, como los Masked Autoencoders (MAE) (He et al., 2022), que entrenan un encoder-decoder para reconstruir parches ocultos en espacio de píxeles.

Sin embargo, reconstruir en píxel puede inducir un sesgo hacia detalles de baja relevancia semántica. Como alternativa, las arquitecturas predictivas en espacio latente —como las Joint-Embedding Predictive Architectures (JEPA)— plantean un objetivo de alineamiento entre representaciones de regiones visibles y regiones ocultas, evitando reconstrucción explícita de textura fina (Assran et al., 2023). Este enfoque promueve invariancia a perturbaciones superficiales y favorece la captura de regularidades estructurales.

En regímenes de pocos datos, el diseño del tokenizador visual resulta determinante. Tokenizadores convolucionales dispersos introducen sesgos inductivos locales que estabilizan el entrenamiento bajo enmascarado y reducen la dependencia de grandes volúmenes de datos (García et al., 2025a). Estas propiedades son especialmente relevantes cuando la representación visual se empleará posteriormente como espacio de optimización para planificación.

Desde la perspectiva de este trabajo, la representación auto-supervisada no es un fin en sí mismo, sino un componente estructural que debe satisfacer un requisito adicional: producir un espacio latente geoméricamente adecuado para descenso energético. Esto implica que pequeñas variaciones en el embedding correspondan, en la medida de lo posible, a transformaciones semánticamente coherentes del entorno. En este sentido,

las arquitecturas predictivas en espacio latente constituyen una base conceptualmente alineada con planificación energética, al priorizar estructura sobre reconstrucción superficial.

3. Método propuesto: Planificación visual por descenso de gradiente

3.1. Planteamiento del problema

Consideramos un sistema dinámico parcialmente observable descrito por estados ocultos $s_t \in \mathcal{S}$ y observaciones visuales $I_t \in \mathcal{I}$. La dinámica real evoluciona según $s_{t+1} \sim p(s_{t+1} | s_t, a_t)$, mientras que el agente únicamente dispone de observaciones generadas por un proceso perceptivo $I_t \sim p(I_t | s_t)$.

El objetivo consiste en alcanzar una condición g , que puede expresarse como un estado objetivo, una configuración visual deseada o una representación embebida. Formalmente, buscamos una secuencia de acciones que induzca una trayectoria coherente con dicho objetivo.

En lugar de optimizar directamente en el espacio de estados \mathcal{S} o en el espacio de observaciones \mathcal{I} , proponemos operar en un espacio latente $\mathcal{Z} \subset \mathbb{R}^d$ inducido por un codificador visual. Definimos la aplicación:

$$f_\phi : \mathcal{I} \rightarrow \mathcal{Z},$$

tal que $z_t = f_\phi(I_t)$ preserve la estructura relevante para planificación. El problema se reformula entonces como la búsqueda de una trayectoria latente $\mathbf{z}_{t:t+H}$ que minimice una energía condicionada al contexto pasado y al objetivo.

3.2. Representación visual semántica

Definimos un codificador visual f_ϕ entrenado de forma auto-supervisada:

$$z_t = f_\phi(I_t), \quad z_t \in \mathbb{R}^d. \quad (1)$$

El propósito del embedding no es únicamente reducir dimensionalidad, sino capturar factores estructurales relevantes para la tarea. Idealmente, variaciones en z_t deben corresponder a transformaciones coherentes del entorno físico, reduciendo sensibilidad a cambios fotométricos o de textura irrelevantes para la planificación.

Como instancia concreta, consideramos una arquitectura compuesta por: (i) un tokenizador convolucional disperso que introduce sesgos inductivos locales y estabilidad en regímenes de pocos datos (García et al., 2025a), y (ii) un objetivo predictivo enmascarado en espacio latente inspirado en JEPa (Assran et al., 2023). Este esquema evita reconstrucción explícita en píxel y promueve invariancia frente a variaciones superficiales.

No se imponen restricciones geométricas explícitas adicionales sobre \mathcal{Z} . La regularidad del espacio latente emerge del objetivo auto-supervisado y de la distribución de datos observados, proporcionando una base estructurada para procesos posteriores de optimización.

3.3. Modelo energético

Definimos una función energética diferenciable:

$$E_\theta(\mathbf{z}_{t:t+H} | \mathbf{z}_{t-K:t}, g, \lambda), \quad (2)$$

donde:

- $\mathbf{z}_{t-K:t}$ representa el contexto pasado,
- g codifica el objetivo,
- λ parametriza preferencias temporales o perfiles de convergencia.

La energía actúa como un funcional escalar sobre trayectorias latentes completas. Su papel es inducir un campo de gradiente que guíe la optimización hacia trayectorias coherentes con la dinámica observada y el objetivo especificado.

El entrenamiento del modelo energético se realiza exactamente bajo el mismo esquema que en Planning-as-Descent (García et al., 2025b). En particular, se ajusta para asignar menor energía a trayectorias observadas en datos offline que conducen al objetivo, y mayor energía a trayectorias inconsistentes, preservando la misma formulación de pérdida y el mismo procedimiento de optimización.

En esta extensión visual no se modifica el objetivo de entrenamiento del modelo energético; la única diferencia reside en que las trayectorias ahora se expresan en el espacio latente visual \mathcal{Z} en lugar de en un espacio de estados explícito.

3.4. Inferencia como descenso de gradiente

La inferencia se realiza mediante refinamiento iterativo de una trayectoria inicial $\mathbf{z}^{(0)}$:

$$\mathbf{z}^{(i+1)} = \mathbf{z}^{(i)} - \eta \nabla_{\mathbf{z}} E_\theta(\mathbf{z}^{(i)} | \mathbf{z}_{t-K:t}, g, \lambda), \quad (3)$$

donde η es el paso de descenso y $i = 0, \dots, N-1$.

Este procedimiento implementa un flujo de gradiente discreto sobre el espacio latente. La dirección de actualización está determinada exclusivamente por la geometría inducida por E_θ . No se introduce ninguna interpretación probabilística ni mecanismos adicionales de proyección: la coherencia dinámica emerge de la forma de la energía aprendida.

La optimización se realiza sobre la trayectoria completa de forma paralela, evitando acumulación autoregresiva de error y facilitando esquemas de replanteamiento frecuente.

3.5. De trayectorias latentes a acciones

Una vez obtenida la trayectoria latente refinada $\hat{\mathbf{z}}_{t:t+H}$, la ejecución puede implementarse mediante:

- **Modelo inverso:** aprender una política $\pi_\psi(a_t | z_t, z_{t+1})$ que aproxime la transición latente planificada.
- **Replanteamiento continuo:** ejecutar el primer paso y reoptimizar tras cada nueva observación.

Esta modularidad separa claramente percepción, planificación y control, permitiendo adaptar cada componente de forma independiente.

3.6. Limitaciones de la planificación en píxeles

Planificar directamente en espacio de píxeles presenta limitaciones estructurales más allá de la dimensionalidad:

- El espacio de píxeles no está alineado con la dinámica física subyacente; pequeñas variaciones en el estado pueden inducir transformaciones complejas en apariencia.
- Los gradientes energéticos pueden alinearse con variaciones fotométricas en lugar de transformaciones relevantes para la tarea.
- El paisaje de optimización en píxel tiende a ser más irregular y sensible a ruido.

En contraste, un espacio latente semántico proporciona una representación más estructurada donde las direcciones de descenso energético tienden a corresponder con transformaciones funcionales del entorno. La hipótesis central es que esta alineación estructural mejora la estabilidad y eficacia del proceso de planificación por descenso.

4. Protocolo experimental

El objetivo experimental es evaluar si la extensión visual propuesta preserva las propiedades de planificación observadas en el espacio de estados cuando se opera en un espacio latente visual auto-supervisado.

Dado que el modelo energético y su entrenamiento son idénticos a los descritos en Planning-as-Descent (García et al., 2025b), el protocolo se centra exclusivamente en analizar el impacto del modelo de observación.

4.1. Entorno y datos

Proponemos utilizar **OGBench** en la tarea **visual-cube-single** como banco de pruebas principal para la versión visual del método. En particular, se consideran dos variantes del entorno:

- **visual-cube-single-play**: conjunto offline de trayectorias de tipo *play* a partir de una política experta de enlace abierto y no markoviana con ruido correlacionado temporalmente.
- **visual-cube-single-noisy**: variante a partir de una política de enlace cerrado con mayor ruido Gaussiano.

El conjunto de datos offline se compone de episodios $\tau = \{(I_t, a_t)\}_{t=1}^T$ disponibles en OGBench, donde I_t son observaciones RGB.

Para garantizar comparabilidad con el trabajo previo, se utilizarán también las versiones equivalentes en espacio de estado reportadas en Planning-as-Descent (García et al., 2025b). De este modo, el rendimiento de **Visual-PaD** podrá contrastarse directamente con las métricas de referencia obtenidas en el espacio de estados s_t bajo el mismo entorno, difiriendo únicamente el modelo de observación.

4.2. Entrenamiento

El entrenamiento se divide en dos etapas:

1) *Preentrenamiento visual*. El codificador f_ϕ se entrena mediante modelado predictivo enmascarado en espacio latente siguiendo (García et al., 2025a; Assran et al., 2023). Esta etapa utiliza únicamente observaciones visuales, sin acceso a acciones ni recompensas.

2) *Entrenamiento del modelo energético*. Una vez fijado el encoder, las trayectorias latentes $z_t = f_\phi(I_t)$ sustituyen a los estados en el procedimiento original de PaD. El modelo energético se entrena sin modificaciones adicionales respecto a (García et al., 2025b).

No se realiza ajuste conjunto inicial entre encoder y energía, con el fin de aislar el efecto del espacio latente aprendido.

4.3. Variantes comparativas

Se comparan tres configuraciones, manteniendo idénticos el horizonte H , contexto K y el presupuesto de inferencia (N iteraciones, paso η):

- **PaD-Estado (referencia)**: planificación original en espacio de estados explícito s_t , utilizando las configuraciones y resultados reportados en (García et al., 2025b).
- **PaD-Píxel (baseline naïve)**: planificación en embeddings aprendidos mediante pre-entrenamiento tipo MAE que reconstruye en espacio de píxel.
- **Visual-PaD (propuesto)**: planificación en embeddings $z_t = f_\phi(I_t)$, con f_ϕ entrenado mediante auto-supervisión sobre los datos de **single-cube-visual-play** y/o **single-cube-visual-noisy**.

El diseño experimental permite aislar el efecto del espacio latente visual: la comparación PaD-Estado vs. Visual-PaD evalúa la pérdida (o preservación) de capacidad de planificación al reemplazar el estado por percepción; mientras que Visual-PaD vs. PaD-Píxel evalúa el beneficio de introducir semántica visual auto-supervisada frente a al espacio de píxel.

La comparación con PaD-Estado se plantea tomando como referencia los resultados ya reportados en (García et al., 2025b), mientras que el presente trabajo define la integración metodológica y el protocolo necesarios para reproducir dicha evaluación en el caso visual.

4.4. Métricas

Se evalúan las siguientes métricas:

- **Tasa de éxito**: porcentaje de episodios que alcanzan el objetivo dentro del horizonte máximo.
- **Longitud efectiva del plan**: número medio de pasos hasta éxito.

Todas las métricas se reportarán promediadas sobre múltiples semillas independientes.

4.5. Hipótesis experimentales

El protocolo permite contrastar las siguientes hipótesis:

- **H1:** La planificación en espacio latente visual aproxima el rendimiento observado en espacio de estados cuando el embedding captura factores relevantes.
- **H2:** La planificación directa en píxel presenta menor estabilidad del descenso y mayor sensibilidad a perturbaciones visuales.
- **H3:** La calidad estructural del embedding influye directamente en la suavidad del paisaje energético y en la eficiencia de convergencia.

La validación empírica de estas hipótesis constituye trabajo en curso.

4.6. Ablaciones sugeridas

Con el fin de analizar el impacto de cada componente en el rendimiento final, se proponen las siguientes ablaciones:

- **Congelar vs. ajustar finamente f_ϕ :** evaluar si el espacio latente aprendido de forma auto-supervisada es suficientemente estructurado para planificación sin adaptación adicional, o si el ajuste conjunto con E_θ mejora alineación entre representación y optimización energética.
- **Tipo de objetivo:** comparar objetivos definidos en espacio de estado, imagen objetivo I^* o embedding z^* , para analizar cómo la modalidad del objetivo afecta la estabilidad del descenso y la eficiencia del plan.
- **Horizonte H e iteraciones N :** estudiar la relación entre longitud de planificación y número de pasos de refinamiento, analizando posibles fenómenos de inestabilidad o saturación del descenso.
- **Frecuencia de replanteamiento (k pasos):** evaluar el compromiso entre coste computacional y robustez dinámica bajo esquemas de planificación tipo MPC en espacio latente.

Estas ablaciones permiten aislar el efecto de la representación visual, la parametrización temporal y la interacción entre percepción y planificación.

5. Discusión y trabajo en curso

La propuesta presentada preserva el núcleo conceptual de Planning-as-Descent —modelo energético y refinamiento por gradiente— y modifica exclusivamente el modelo de observación para admitir entradas visuales. Esta separación explícita permite analizar de forma controlada el impacto del espacio de representación sobre el proceso de optimización.

Desde un punto de vista estructural, el enfoque introduce tres implicaciones relevantes:

- **Desacoplamiento percepción–planificación:** el encoder visual se entrena sin señales de control, lo que favorece reutilización y modularidad.

- **Regularización implícita del espacio de optimización:** la auto-supervisión induce un embedding estructurado que puede facilitar descensos más estables frente a optimización directa en píxel.
- **Escalabilidad perceptiva:** la extensión a visión permite aplicar el paradigma energético a escenarios donde el estado no es directamente observable.

No obstante, la propuesta introduce interrogantes abiertos. En particular, no está garantizado que el embedding visual preserve información dinámica suficiente para planificación precisa. Asimismo, la alineación entre estructura semántica y dinámica física puede depender críticamente del tipo de datos utilizados durante el preentrenamiento.

El trabajo en curso aborda la validación empírica del protocolo definido en OGBench visual-cube-single, así como el análisis comparativo con las versiones en espacio de estado reportadas en (García et al., 2025b). Además, se exploran extensiones hacia:

- ajuste conjunto encoder–energía,
- transferencia a entorno real,
- incorporación de sensores adicionales (p.ej., propiocepción del robot).

En conjunto, este trabajo constituye un paso intermedio hacia una formulación unificada donde representación auto-supervisada y planificación energética coexistan en entornos parcialmente observables.

Agradecimientos

Este trabajo ha sido realizado parcialmente gracias al apoyo del Centro para el Desarrollo Tecnológico y la Innovación (CD-TI) – Ministerio de Ciencia, Innovación y Universidades, en el marco del Programa CERVERA, en el proyecto ROCCA bajo el expediente CER 20251010.

Referencias

- Assran, M., Duval, Q., Misra, I., Bojanowski, P., Vincent, P., Rabat, M., LeCun, Y., Ballas, N., 2023. Self-supervised learning from images with a joint-embedding predictive architecture. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. pp. 15619–15629.
- Carroll, M., Paradise, O., Lin, J., Georgescu, R., Sun, M., Bignell, D., Milani, S., Hofmann, K., Hausknecht, M., Dragan, A., et al., 2022. Uni [mask]: Unified inference in sequential decision problems. Advances in neural information processing systems 35, 35365–35378.
- Chen, L., Lu, K., Rajeswaran, A., Lee, K., Grover, A., Laskin, M., Abbeel, P., Srinivas, A., Mordatch, I., 2021. Decision transformer: Reinforcement learning via sequence modeling. Advances in neural information processing systems 34, 15084–15097.
- Chi, C., Xu, Z., Feng, S., Cousineau, E., Du, Y., Burchfiel, B., Tedrake, R., Song, S., 2025. Diffusion policy: Visuomotor policy learning via action diffusion. The International Journal of Robotics Research 44 (10-11), 1684–1704.
- Florence, P., Lynch, C., Zeng, A., Ramirez, O. A., Wahid, A., Downs, L., Wong, A., Lee, J., Mordatch, I., Tompson, J., 2022. Implicit behavioral cloning. In: Conference on robot learning. PMLR, pp. 158–168.
- García, C. V., Cazoria, M., Pomares, J., 2025a. Escaping the big data paradigm in self-supervised representation learning. arXiv preprint arXiv:2502.18056 Accepted for publication at Computer Vision and Image Understanding (CVIU) journal in 2026.

- García, C. V., Cazorla, M., Pomares, J., 2025b. Planning as descent: Goal-conditioned latent trajectory synthesis in learned energy landscapes. arXiv preprint arXiv:2512.17846 Under review at Engineering Applications of Artificial Intelligence (EAAD) journal.
- He, K., Chen, X., Xie, S., Li, Y., Dollár, P., Girshick, R., 2022. Masked autoencoders are scalable vision learners. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. pp. 16000–16009.
- Henaff, M., Canziani, A., LeCun, Y., 2019. Model-predictive policy learning with uncertainty regularization for driving in dense traffic. arXiv preprint arXiv:1901.02705.
- Janner, M., Du, Y., Tenenbaum, J. B., Levine, S., 2022. Planning with diffusion for flexible behavior synthesis. arXiv preprint arXiv:2205.09991.
- LeCun, Y., et al., 2022. A path towards autonomous machine intelligence version 0.9. 2, 2022-06-27. Open Review 62 (1), 1–62.
- Wu, P., Majumdar, A., Stone, K., Lin, Y., Mordatch, I., Abbeel, P., Rajeswaran, A., 2023. Masked trajectory models for prediction, representation, and control. In: International Conference on Machine Learning. PMLR, pp. 37607–37623.