Qeios

Peer Review

Review of: "Unified Guidance for Geometry-Conditioned Molecular Generation"

Emmanuel Salawu[1]

1. Google (United States), Mountain View, United States

The manuscript introduces UniGuide, a novel framework designed to unify geometric guidance for molecular generation using diffusion models. The core idea is to use **condition maps** to bridge arbitrary source conditions with target conditions suitable for guiding unconditional diffusion models. This approach aims to address limitations in current molecular diffusion models, which are often tailored to specific downstream tasks and, thus, lack adaptability. UniGuide distinguishes itself by not requiring additional training or external networks to guide the generation process.

The manuscript presents a valuable contribution to molecular generation. UniGuide offers a unified and flexible framework for geometry-conditioned drug design. While the manuscript has some limitations, its strengths outweigh its weaknesses.

Strengths

- **Unified Approach:** The manuscript's strength lies in its unified approach to geometry-conditioned molecular generation. By using condition maps, UniGuide can handle structure-based, fragment-based, and ligand-based drug design tasks within a single framework. This is a significant advantage over specialized models that lack adaptability.
- **Flexibility and Adaptability:** UniGuide's flexible formulation can be generalized to new geometric tasks, such as conditioning on atomic densities. The separation of model training and conditioning allows it to tackle tasks even with minimal data, which is crucial in the biological domain.
- **Performance:** UniGuide demonstrates competitive or superior performance compared to task-specific baselines. For instance, in ligand-based drug design (LBDD), UniGuide achieves higher shape similarity than ShapeMol+g, even though the latter uses position correction techniques.

doi.org/10.32388/795MYD

- **Focus on Unconditional Generation:** By redirecting focus to advancing unconditional generation, UniGuide benefits multiple applications and allows tackling tasks with limited data. This is a notable shift, as much of the novelty in conditional models comes from condition incorporation.
- **Comprehensive Evaluation:** The manuscript evaluates UniGuide on a variety of drug discovery tasks, including ligand-based (LBDD), structure-based (SBDD), and fragment-based drug design (FBDD). The experiments demonstrate UniGuide's practical relevance and transferability to diverse scenarios.

Weaknesses and Areas for Improvement

- **Limited Scope:** UniGuide excludes tasks beyond purely geometric conditions, such as those encompassing global graph properties. The manuscript should acknowledge this limitation more explicitly.
- **Dependency on Configuration Space:** UniGuide requires the unconditional model to be trained on a matching configuration space. This requirement may limit its applicability in certain scenarios.
- **Runtime Efficiency:** UniGuide has a higher runtime compared to other conditioning mechanisms because it computes gradients through the diffusion model at inference time. This should be addressed by discussing potential optimization strategies.
- **Lack of Guarantee for Fragment Presence:** For fragment-based drug design (FBDD), UniGuide cannot guarantee that the condition fragments are present in the generated samples. The manuscript mentions a post-hoc step to address this, but further investigation is needed.
- **Reproducibility and Implementation Details:** While the manuscript provides implementation details in the appendices, it should ensure that all necessary information is readily accessible for reproducibility. This includes details on datasets, training parameters, and code availability.

Additional Comments and Questions

- **Condition Map Derivation:** The manuscript mentions deriving a condition map for ligand-based drug design (LBDD). It would be helpful to provide more details on the general principles for deriving such maps for different geometric constraints.
- **Equivariance Proof:** The proof of Theorem 4.1 is mentioned in Appendix B. The manuscript should highlight the key steps and assumptions in the main text to improve readability.
- **Comparison with Validity Guidance:** The comparison of UniGuide with validity guidance in Appendix D.4 is insightful. The manuscript should emphasize the advantages of UniGuide's

separation of surface computation from gradient computation.

- **Impact Statement:** The impact statement in Appendix A acknowledges the potential hazards of generating dangerous substances. The manuscript should elaborate on the safeguards and ethical considerations for using UniGuide in drug discovery.

Declarations

Potential competing interests: No potential competing interests to declare.