Generative Biology

This manuscript ([permalink](#)) was automatically generated from [In-Vivo-Group/generative-biology@8de33bc](#) on November 7, 2023.

Authors

- **Alexander J. Titus** ✉
  ⓘ [0000-0002-0145-9564](#) · ⓞ [alexandertitus](#)
  In Vivo Group, Washington, DC, USA; International Computer Science Institute, Berkeley, CA, USA · Funded by Grant TBD

- **Matthew E. Walsh**
  ⓘ [0000-0003-1514-7761](#) · ⓞ [mwalsh52](#)
  Center for Health Security, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, USA; Department of Environmental Health and Engineering, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, USA

✉ — Correspondence possible via [GitHub Issues](#) or email to Alexander J. Titus <publications@theinvivogroup.com>.

Abstract

The rapid pace of progress in generative artificial intelligence (AI) techniques like deep learning, reinforcement learning, and transformer neural networks is transforming the life sciences and biomedicine. This living review paper provides an updatable, comprehensive overview and analysis of the latest literature on generative biology – the application of cutting-edge generative AI methods to accelerate insights and innovation across the life sciences and healthcare. All authors are welcome to contribute to this review via pull requests.

The review synthesizes key developments in using generative models for de novo biomedical discovery, design, and decision support. It examines techniques and applications including deep learning on omics data for personalized medicine, generative chemistry for drug development, protein structure prediction for molecular engineering, image synthesis for pathology, language models for clinical decision support, robotic simulation for prosthetics, and generative networks for cell programming.

The review highlights representative studies and benchmarks in each area while contextualizing progress, limitations, emerging best practices, and directions for future work. It also discusses social and ethical challenges raised by generative biology applications, such as compounding bias, system opacity, and dual-use risks, alongside proposed solutions.

As a living review, this paper will be continually updated as the field rapidly advances to provide researchers and practitioners with an up-to-date reference on the state of the art in employing generative AI to accelerate biomedicine for the collective good.

Executive Summary

The goal with be a 2 page TL;DR of the review after v1 is complete. Need more content.

Introduction

This is the start of the Generative Biology living review!

Recent Advances in Generative AI

Introduction

Brief background on rise of generative models like GPT-3, DALL-E, AlphaFold, etc. Summary of scope and goals of chapter focused on key advances in last 1-2 years.

Transformer-Based Language Models

GPT-3 and Foundation Models

- Overview of GPT-3 architecture and self-supervised training on massive text corpus
- Discussion of GPT-3 capabilities and limits, including few-shot learning
- Concept of foundation models as basis for many downstream applications

Other Notable Models

- Summary of other major transformer language models like Google's PaLM, DeepMind's Gopher, Meta's OPT, Anthropic's Claude etc.
- Comparison of model sizes, architectures, training approaches
- Benchmarking of models on various NLP tasks

Multimodal Generative Models

DALL-E 2 and Text-to-Image Generation

- Explain DALL-E 2 architecture and training methodology
- Discuss capabilities in text-to-image generation
- Issues around bias, appropriate use cases

Other Multimodal Models

- Overview of models like Imagen, Parti, Flamingo for text-to-image
- Discussion of video generation models like Googles Imagen Video
- Models for text to 3D shapes, text to music etc.

Outlook

- Key challenges and limitations of current generative models

- Likely future advances building on these models
- Broader societal impact of widely available generative models

References