Multi modal learning for drug discovery

Lab: computational bioimaging and bioinformatics, ENS
Location: ENS Paris, 45 rue d'Ulm, 75005 Paris
Contacts: Auguste Genovesio (auguste.genovesio@ens.psl.eu), Ethan Cohen (ecohen@bio.ens.psl.eu)

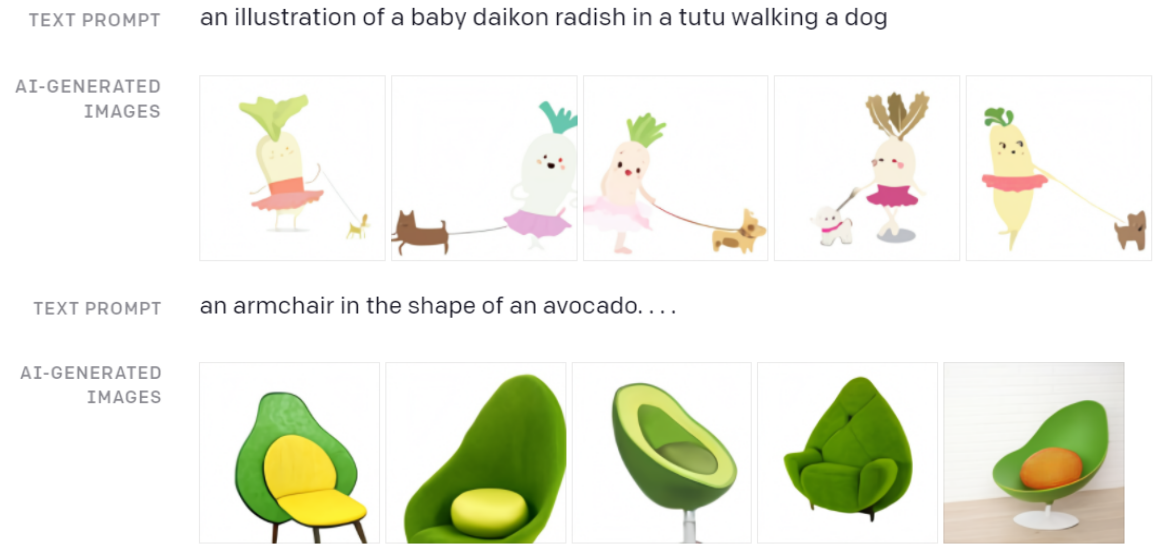

Fig a. Text to image generation with DALLE.

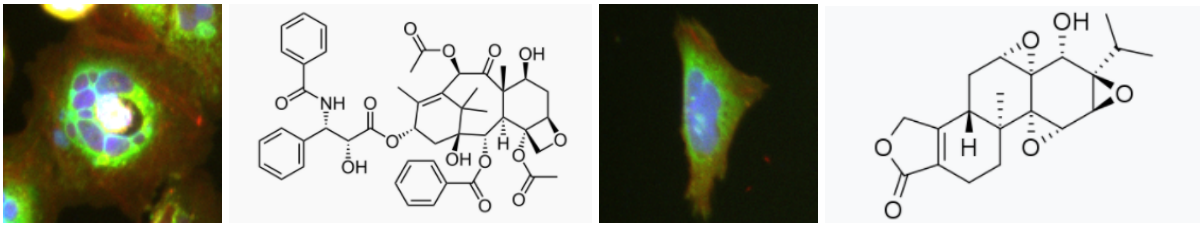

Fig b. Cellular images disturbed by different molecules produce various phenotypes.

Context

Drug discovery is a long process that often takes a decade to design and lead to new therapeutics molecules on the market. The target-based approach that assumes the prior knowledge of a molecular target is currently the main strategy in the pharmaceutical industry. However, an alternative method, the phenotype-based approach has led to many more discoveries of first in class drugs. One of the major tools in this strategy is High Content Screening. By combining robotized experiments and automated microscopy, it allows to image cells under thousands of parallel small molecule perturbations.

Background

In recent years, outstanding improvements have been made in multi-modal learning, especially for text and images data. Methods such as CLIP [1] uses contrastive learning methods to learn mutual representations of text and images. Other methods such as DALLE [2] improved drastically text to image generation tasks, and allows generations of images with high quality, fidelity, and diversity. Such methods have been made possible thanks to recent improvements in Transformers models (both in Language and Vision) as well as multi-modal representation learning.

This project

This project aims at extending such methods to learn meaningful relationships between molecules and cell painting images, as well as producing images of cells based on molecule structures. This idea is based on the fact that one of the common representations of molecules named SMILES, is a textual representation of molecular data.

The intern will:

1. Explore text to image matching/generations methods.
2. Adapt relevant methods to the molecule to cell image generation problem.

Depending on interest, the intern will have the possibility to explore different ways to process molecules (2D graph and 3D).

The candidate:

The candidate should know Python and would ideally have some experience with PyTorch. Experience in multi-modal learning would be considered a plus.

Bibliography

[1] Radford et al, Learning transferable visual models from natural language supervision
[2] Ramesh et al, Zero-Shot Text-to-Image Generation