

# **TTIC 31230, Fundamentals of Deep Learning**

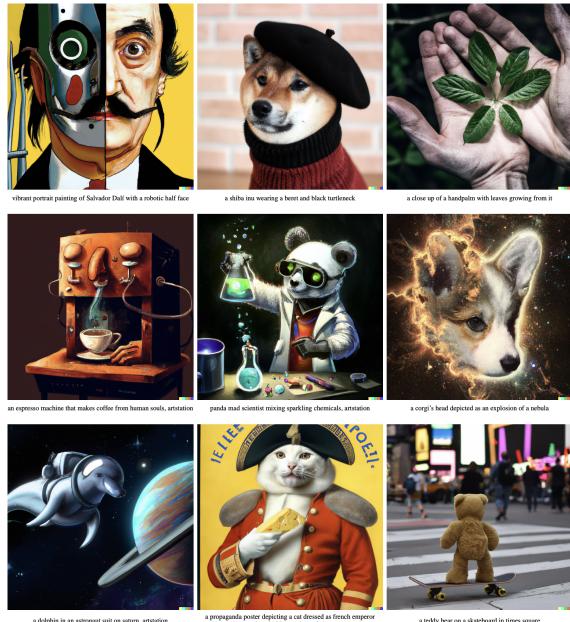
David McAllester, Autumn 2022

DALL·E-2

# DALL·E-2



panda mad scientist mixing sparkling chemicals, artstation



## Conditional Cross Entropy Loss

$$\Phi^* = \operatorname{argmin}_{\Phi} E_{(x,y) \sim \text{Pop}} - \ln P_{\Phi}(y|x)$$

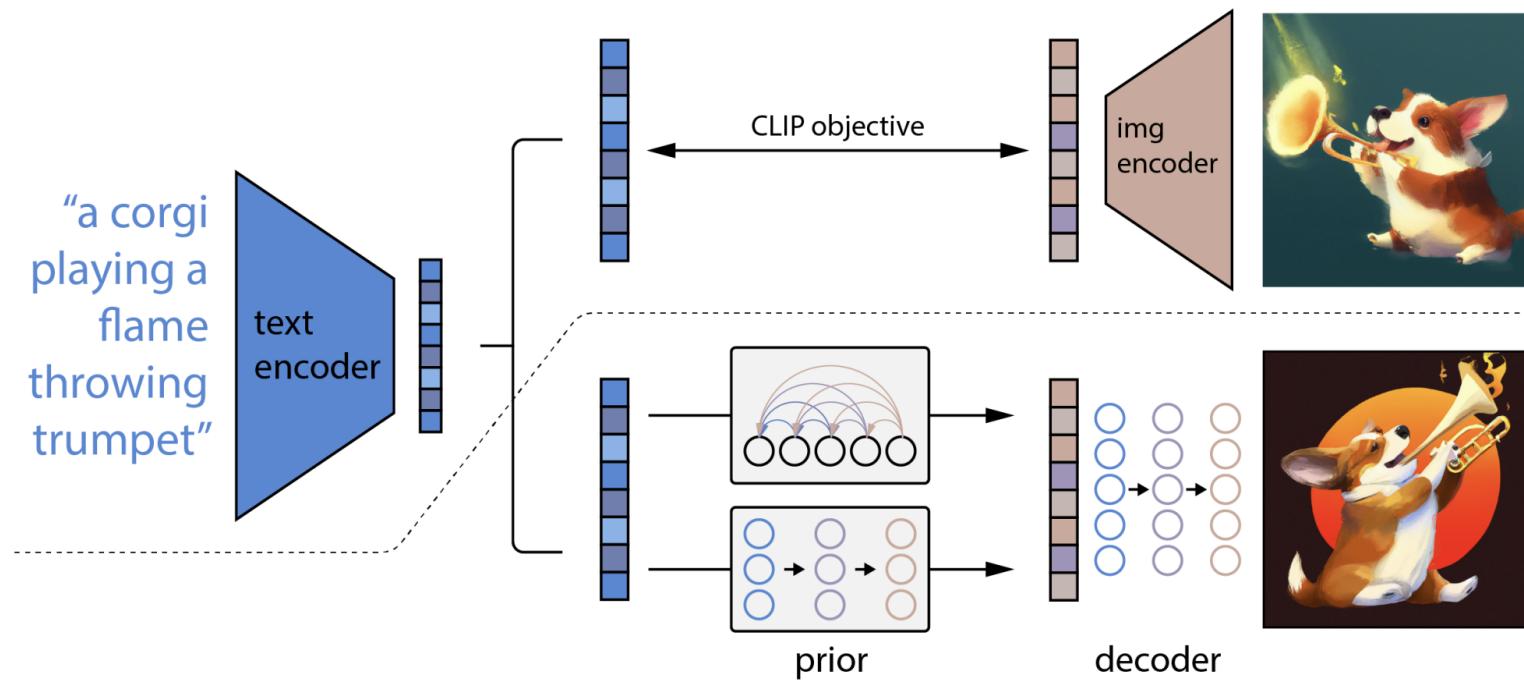
In DALLE we have that  $x$  is text and  $y$  is an image.

We need to formulate **conditional diffusion models** in a way that supports conditioning images on text.

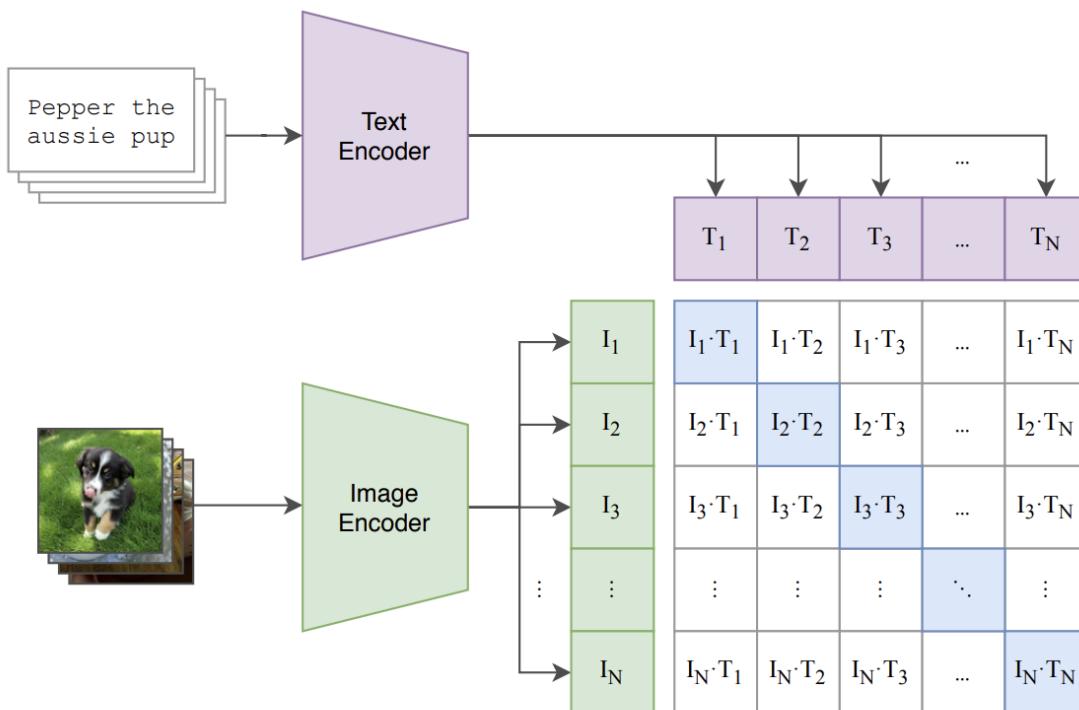
We would like to both be able to compute (or upper bound bound)  $P_{\Phi}(y|x)$  and also to generate  $y$  from  $P_{\Phi}(y|x)$ .

# DALL·E-2

DALLE combines diffusion models with CLIP.



# CLIP Does Contrastive Coding



# GLIDE: Towards Photorealistic Image Generation ...

Nichol, Dhariwal, Ramesh, et al., March 2022

Slides under development.

**END**