Autonomous and Adaptive Systems

Generative Learning

Mirco Musolesi

mircomusolesi@acm.org

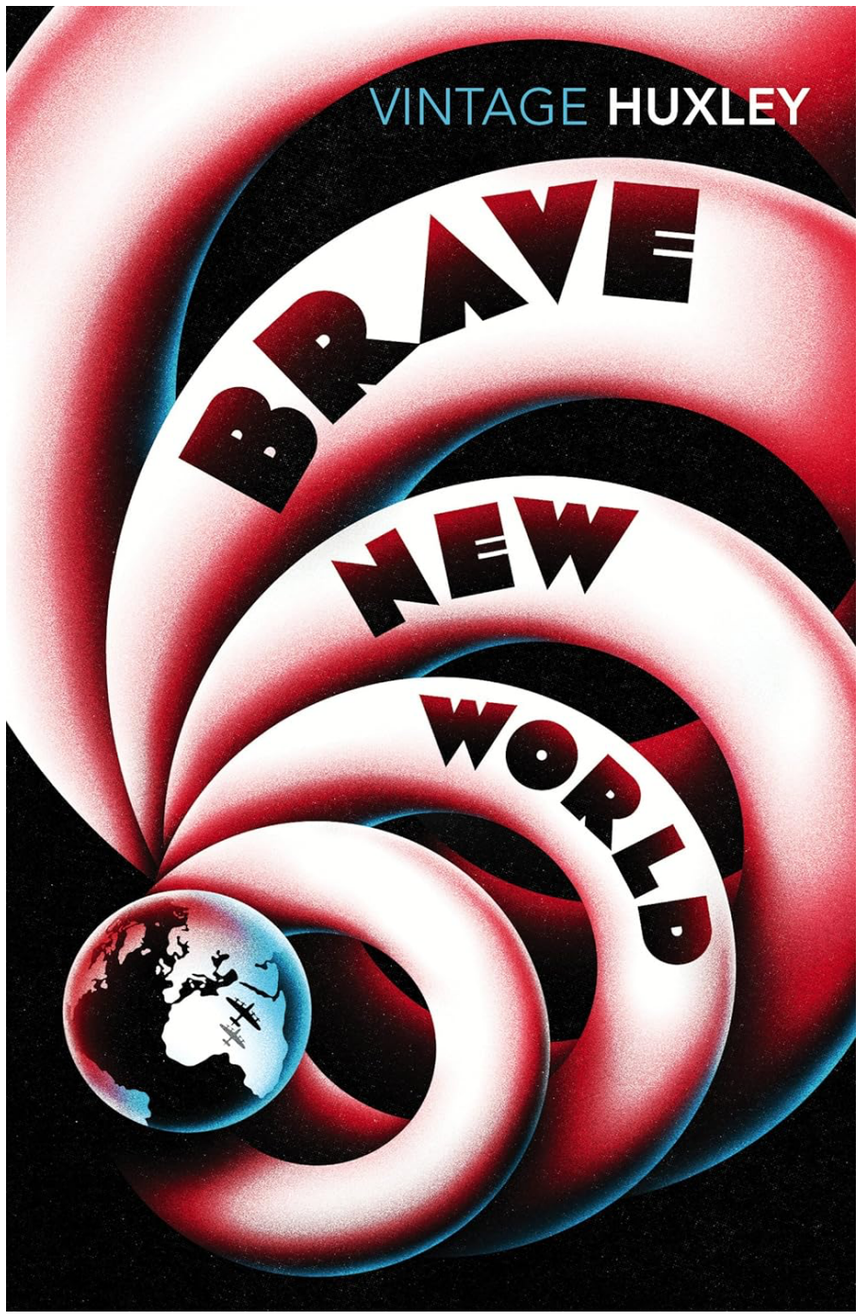
Generative Learning

- In this lecture we will focus on machines that creates and autonomously plan.

- In the recent years, many systems that exhbit the capacity of creating new artefacts and autonomously “invent” new solutions have been presented.

- And we are just at beginning…

A Brave New World



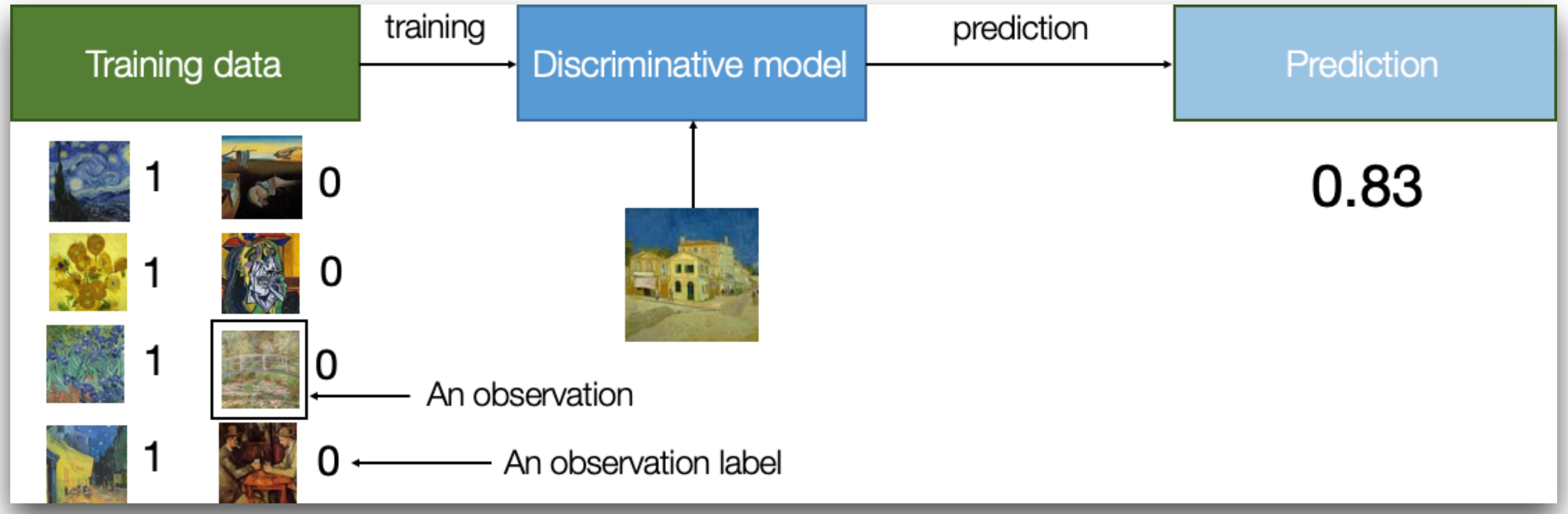
Generative Modeling

- A generative model describes how a dataset is generated for example through a probabilistic description. Through sampling of this model, we generate new data.

- The goal is to generate data are variations of the existing ones, but not “too far” from the original dataset.

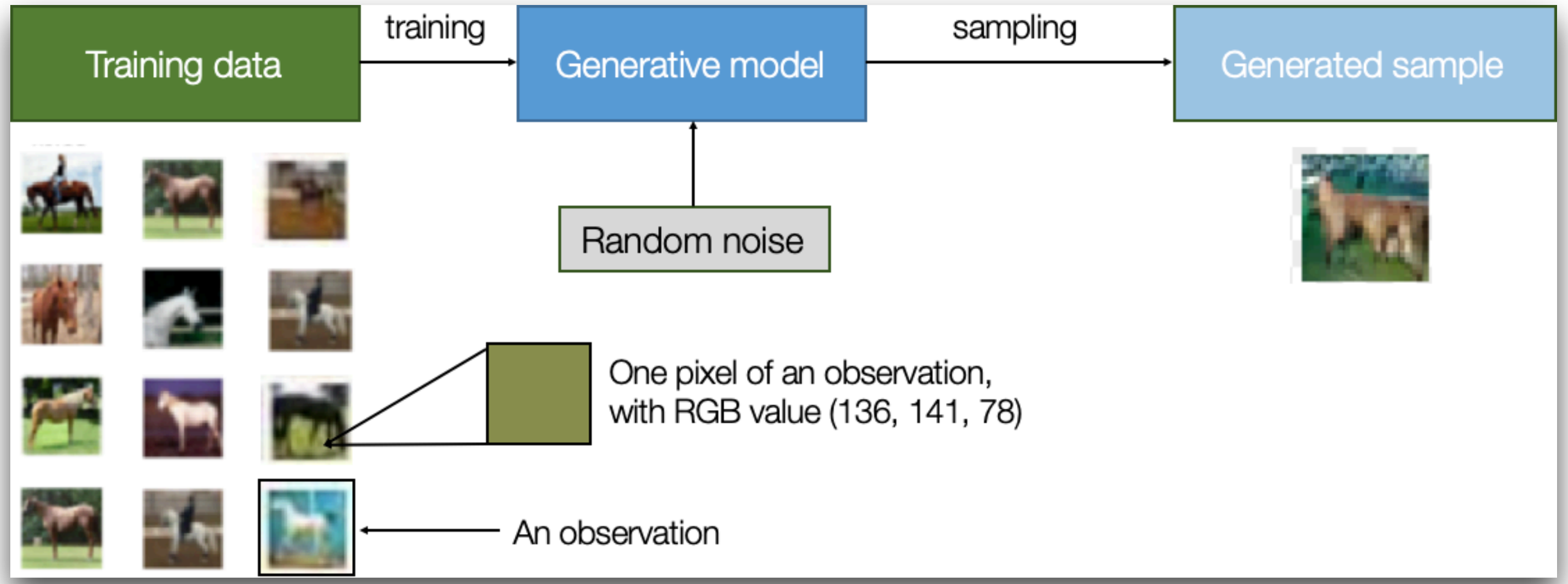
- A generative model is usually probabilistic rather than deterministic in nature.

Discriminative Model



Source: David Foster. Generative Deep Learning. O'Reilly. 2019.

Generative Model



Source: David Foster. Generative Deep Learning. O'Reilly. 2019.

Generative Modelling Framework

- Given a dataset $\mathbf{X}$, we assume that the observation has been generated according to some *unknown* distribution p_{data}.

- The goal is to create a generative model p_{model} that can be used to generate samples that look like they were drawn from p_{data}.

- We achieved our goal if the generated data are also suitably different from the observations in $\mathbf{X}$.

  - The model should not simply reproduce the things that have already been seen.

Variational Autoencoders

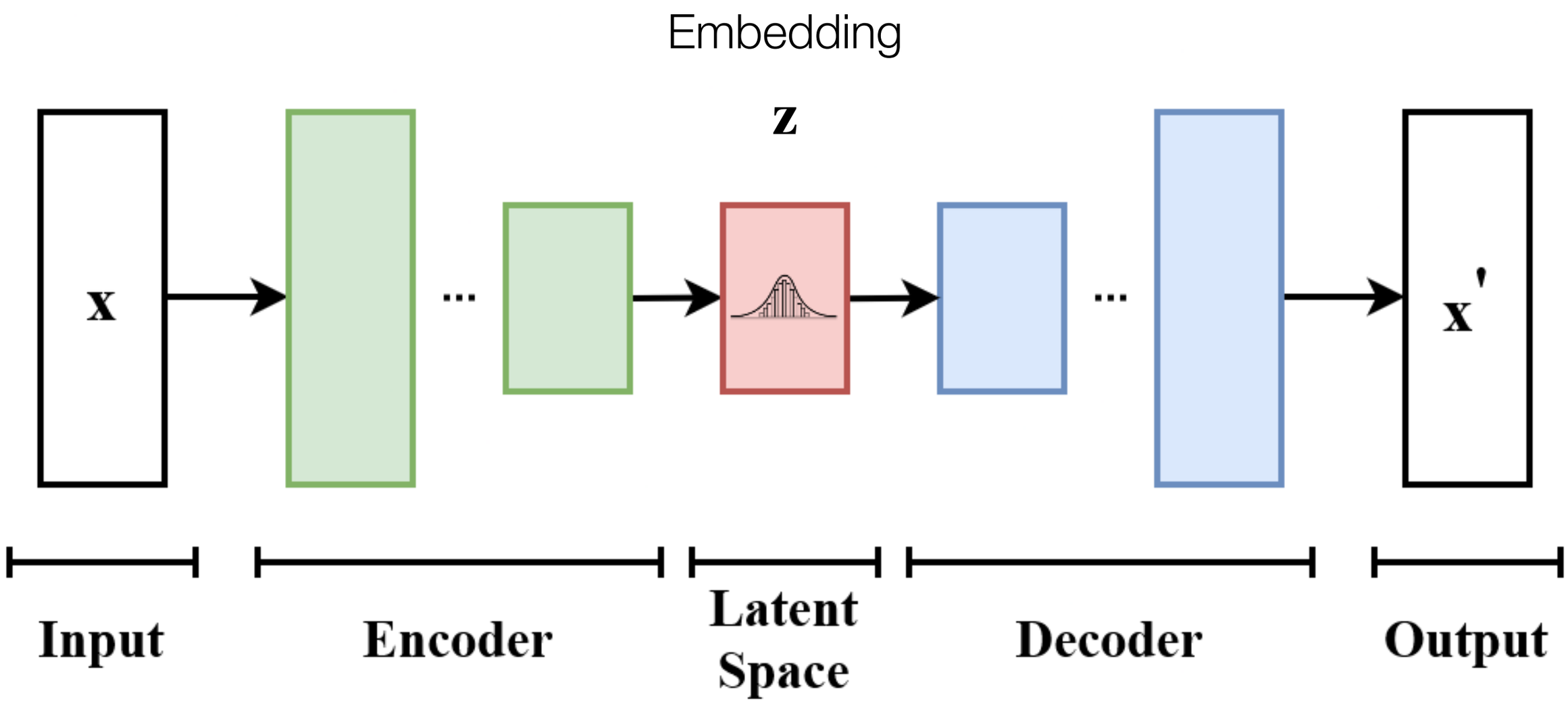
- An *autoencoder* is a neural network that is trained to perform the task of encoding and decoding an item, such that the output from this process is close to the original item as much as possible.

- An autoencoder is composed of two parts:

  - An *encoder network* that compresses high-dimensional input data such an image into a lower-dimensional embedding vector.

  - A *decoder network* that decompresses a given embedding vector back to the original.

Variational Autoencoders



Source: Wikimedia

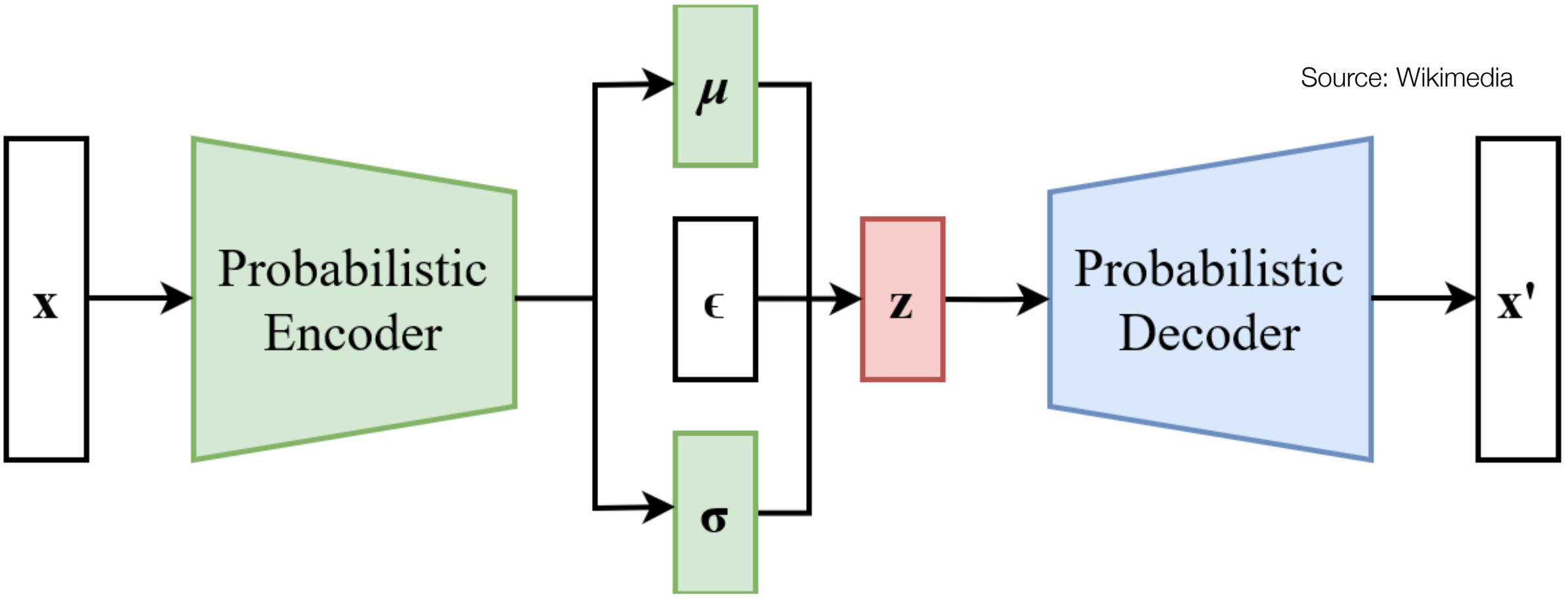
Variational Autoencoders

- The embedding $\mathbf{z}$ is a compression of the original input into a lower-dimensional latent space.

- By sampling the latent space, we can generate new outputs by passing the sample to the decoder, since the decoder has learned how to convert points into “realistic” outputs.

Variational Autoencoders

Source: Wikimedia

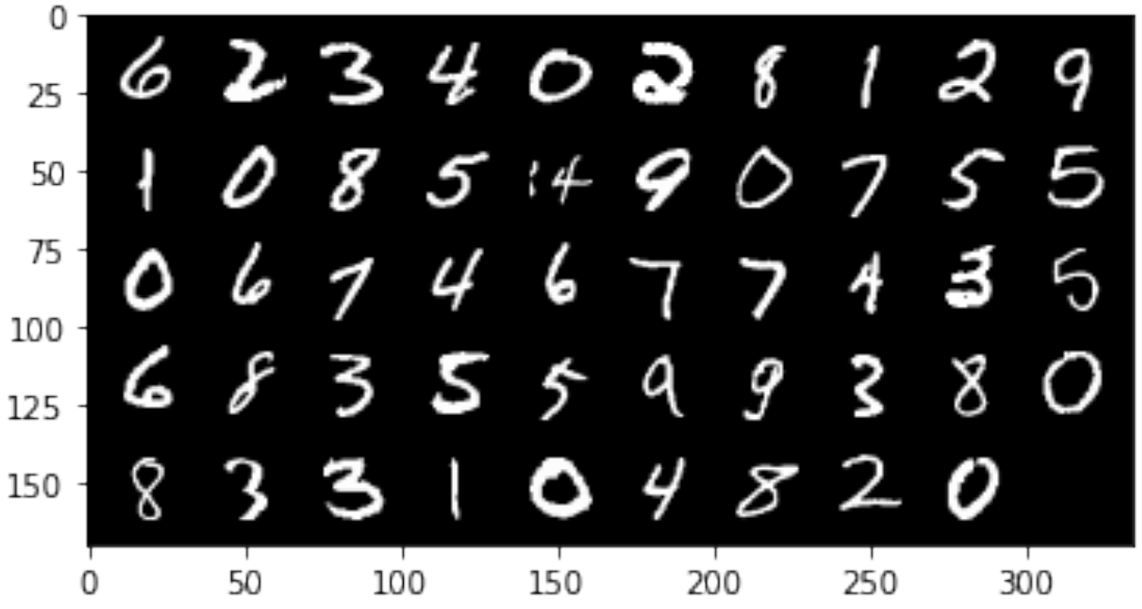


The encoder takes each input image and encode it into two vectors that defines a multivariate normal distribution in the latent space with mean μ and variance σ

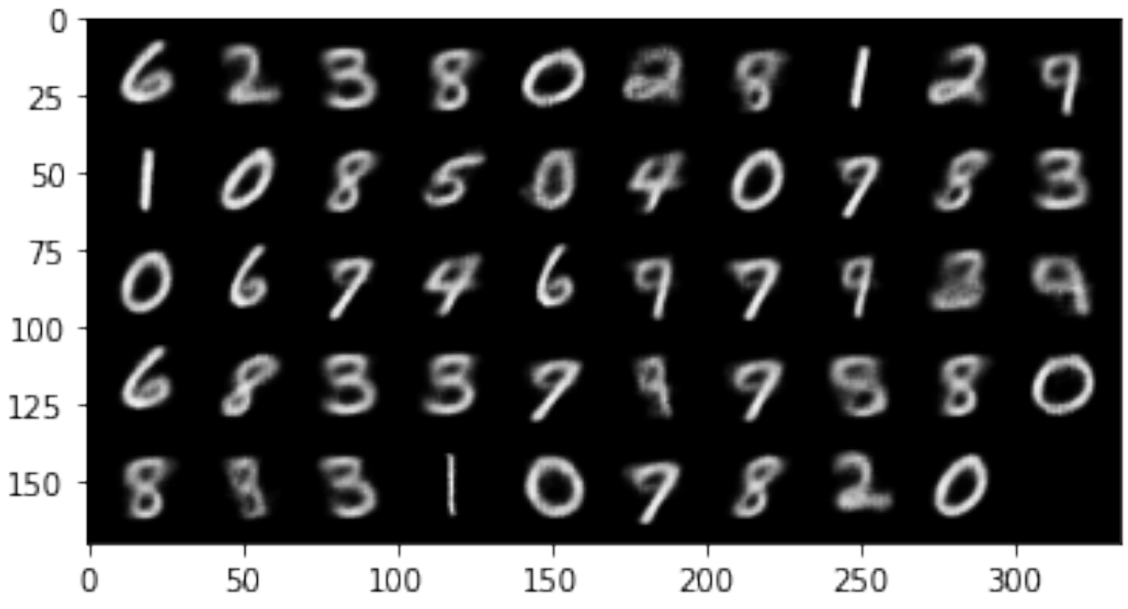
ϵ is sampled from a normal distribution $(\mathbf{0},\mathbf{I})$

We calculate $\mathbf{z}$ as follows $\mathbf{z} = \mu + \sigma\epsilon$

Variational Autoencoders



Original



VAE Reconstruction

Colab notebook: https://colab.research.google.com/github/smartgeometry-ucl/dl4g/blob/master/variational_autoencoder.ipynb

DeepDream

- Developed by Alexander Mordvintsev, Christopher Olah and Mike Tika in 2015.
- This movement is also called inceptionism from Chris Nolan's movie "Inception" (but a bit indirectly).
- The idea is to try to exploit the "patterns" that are learnt by neural network "in reverse".
- This is used to generate images that are composed by the patterns that are detected by the different layers.

Going deeper with convolutions

Christian Szegedy
Google Inc.

Wei Liu
University of North Carolina, Chapel Hill

Yangqing Jia
Google Inc.

Pierre Sermanet
Google Inc.

Scott Reed
University of Michigan

Dragomir Anguelov
Google Inc.

Dumitru Erhan
Google Inc.

Vincent Vanhoucke
Google Inc.

Andrew Rabinovich
Google Inc.

Abstract

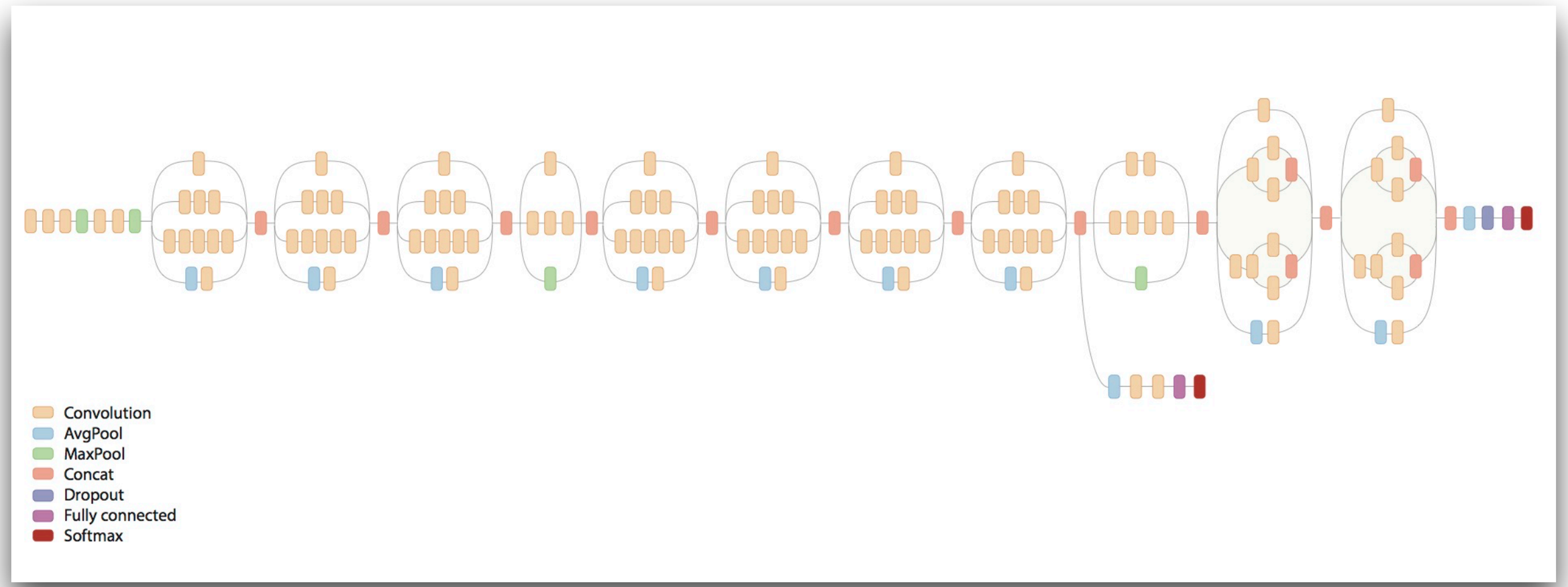
We propose a deep convolutional neural network architecture codenamed Inception, which was responsible for setting the new state of the art for classification and detection in the ImageNet Large-Scale Visual Recognition Challenge 2014 (ILSVRC14). The main hallmark of this architecture is the improved utilization of the computing resources inside the network. This was achieved by a carefully crafted design that allows for increasing the depth and width of the network while keeping the computational budget constant. To optimize quality, the architectural decisions were based on the Hebbian principle and the intuition of multi-scale processing. One particular incarnation used in our submission for ILSVRC14 is called GoogLeNet, a 22 layers deep network, the quality of which is assessed in the context of classification and detection.


WE NEED TO GO
DEEPER

DeepDream

- Interpretability is still an open question in deep learning.
- In case of images, we know that each layer progressively extracts higher and higher-level features of the images, until the final layer makes a decision on what an image actually shows.
  - First layer for edges and corners, then intermediate layers interpret the basic features and are used to extract shapes and components (windows, leaves, etc). A final layer might extract buildings or trees.
- One way to visualise this is to turn the network upside down and ask to enhance an input image so that the role of each layer can be interpreted.

Inception Network



Source: Inception in TensorFlow
https://github.com/tensorflow/models/tree/master/research/inception

Inception Network

Rethinking the Inception Architecture for Computer Vision

Christian Szegedy
Google Inc.
szegedy@google.com

Vincent Vanhoucke
vanhoucke@google.com

Sergey Ioffe
sioffe@google.com

Jonathon Shlens
shlens@google.com

Zbigniew Wojna
University College London
zbigniewwojna@gmail.com

Abstract

Convolutional networks are at the core of most state-of-the-art computer vision solutions for a wide variety of tasks. Since 2014 very deep convolutional networks started to become mainstream, yielding substantial gains in various benchmarks. Although increased model size and computational cost tend to translate to immediate quality gains for most tasks (as long as enough labeled data is provided for training), computational efficiency and low parameter count are still enabling factors for various use cases such as mobile vision and big-data scenarios. Here we are exploring ways to scale up networks in ways that aim at utilizing the added computation as efficiently as possible by suitably factorized convolutions and aggressive regularization. We benchmark our methods on the ILSVRC 2012 classification challenge validation set demonstrate substantial gains over

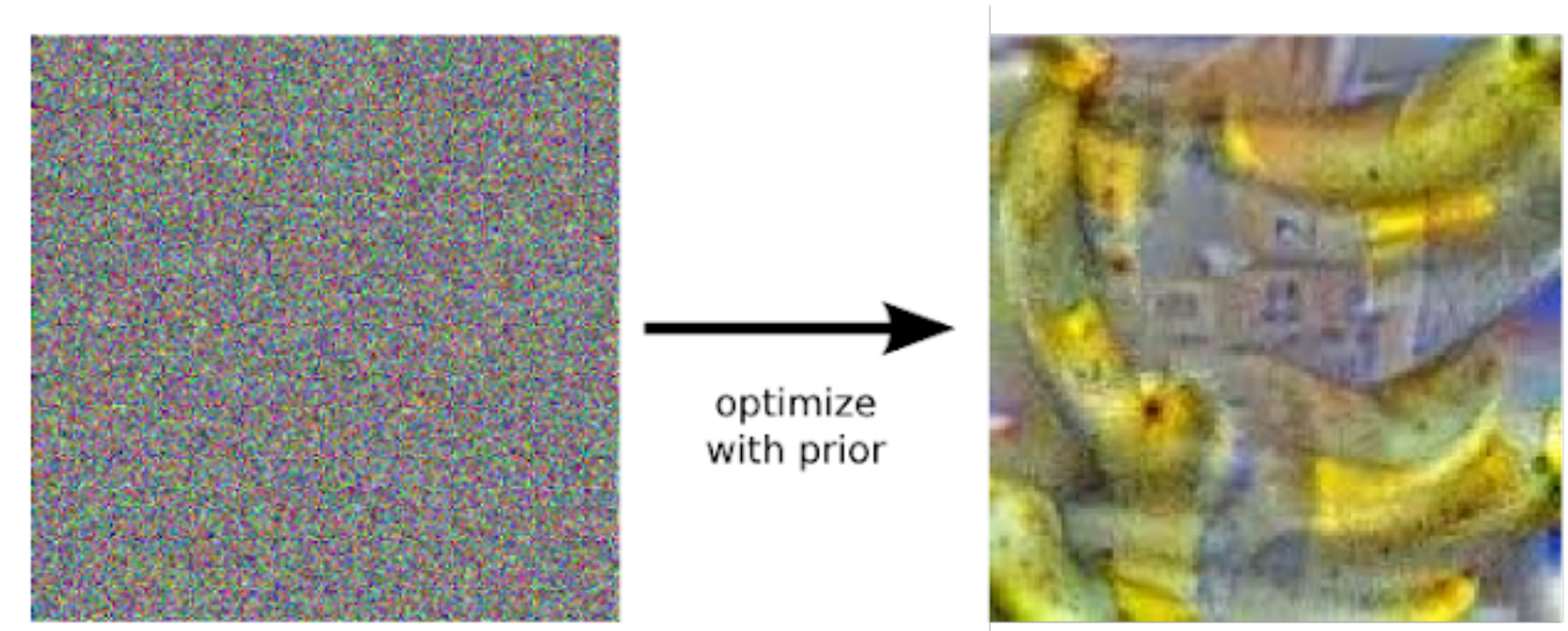
larly high performance in the 2014 ILSVRC [16] classification challenge. One interesting observation was that gains in the classification performance tend to transfer to significant quality gains in a wide variety of application domains. This means that architectural improvements in deep convolutional architecture can be utilized for improving performance for most other computer vision tasks that are increasingly reliant on high quality, learned visual features. Also, improvements in the network quality resulted in new application domains for convolutional networks in cases where AlexNet features could not compete with hand engineered, crafted solutions, e.g. proposal generation in detection[4].

Although VGGNet [18] has the compelling feature of architectural simplicity, this comes at a high cost: evaluating the network requires a lot of computation. On the other hand, the Inception architecture of GoogLeNet [20]

DeepDream

- The idea of DeepDream is to choose a layer (or layers) and minimise the loss in a way that the image increasingly “excites” the layers.
- The complexity of the features incorporated depends on the layer we chose.
- We use the InceptionV3 architecture.
  - For DeepDream the layers of interest are those where the convolutions are concatenated.
- Once we have calculated the loss for the chosen layers, we calculate the gradients (using gradient ascent) to the input images.
- If we consider a “noise” image, the shapes that are identified by that specific layer will appear.

DeepDream

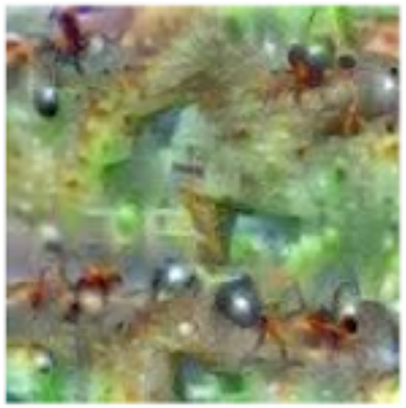


Source: Inceptionism: Going Deeper into Neural Networks
https://ai.googleblog.com/2015/06/inceptionism-going-deeper-into-neural.html

DeepDream

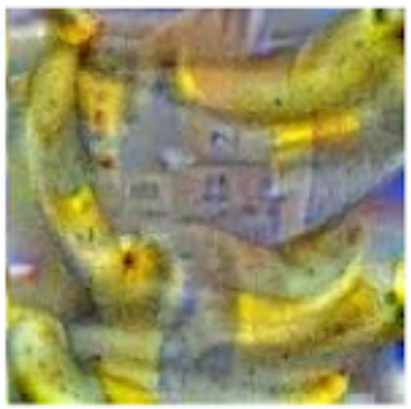

Hartebeest


Measuring Cup


Ant


Starfish


Anemone Fish


Banana


Parachute


Screw

Source: Inceptionism: Going Deeper into Neural Networks
https://ai.googleblog.com/2015/06/inceptionism-going-deeper-into-neural.html

DeepDream

Source: Inceptionism: Going Deeper into Neural Networks
https://ai.googleblog.com/2015/06/inceptionism-going-deeper-into-neural.html

DeepDream

Source: Inceptionism: Going Deeper into Neural Networks
https://ai.googleblog.com/2015/06/inceptionism-going-deeper-into-neural.html

TensorFlow DeepDream Colab

https://colab.research.google.com/github/tensorflow/docs/blob/master/site/en/tutorials/generative/deepdream.ipynb

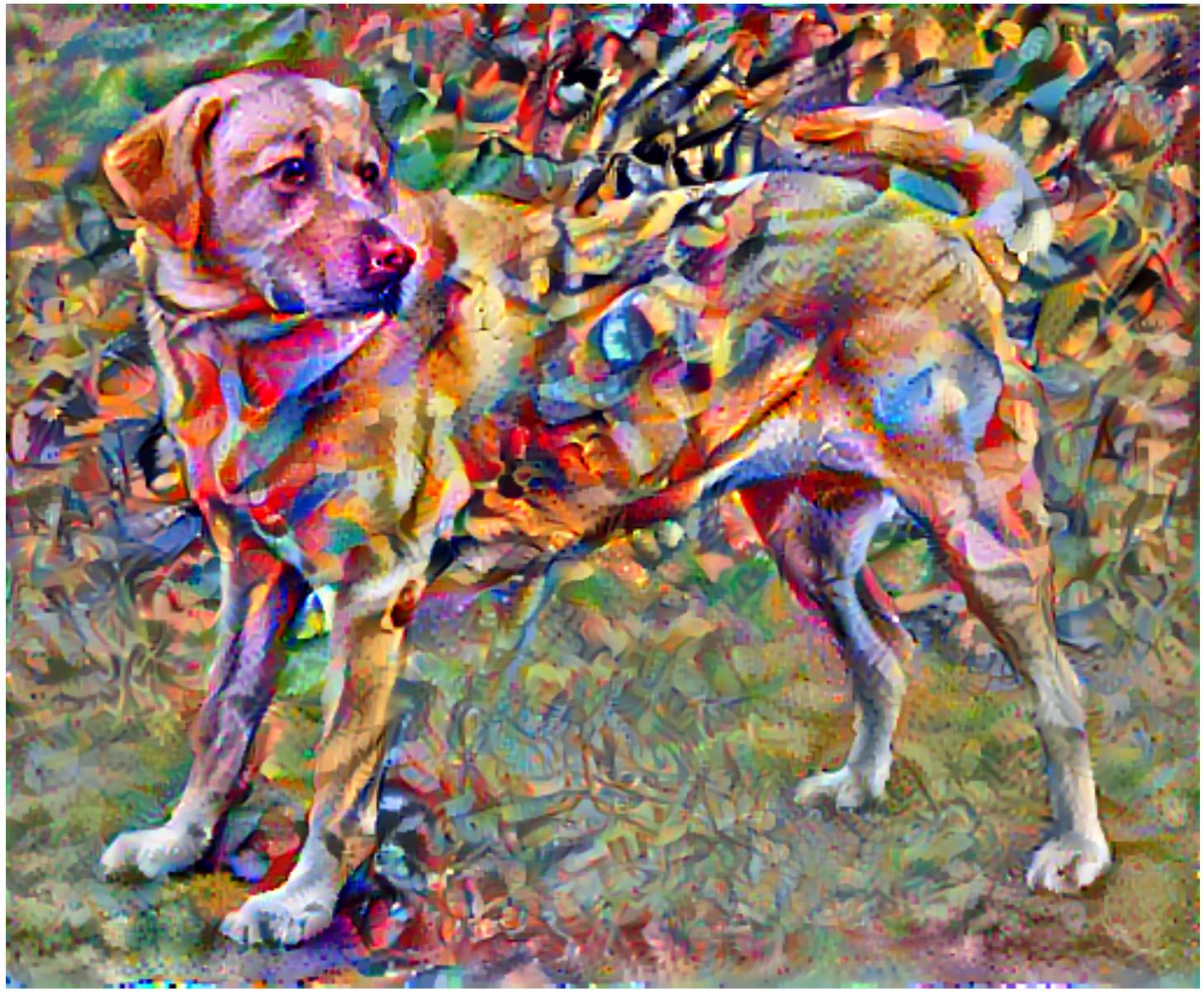
Neural Style Transfer

Source:TensorFlow Neural Style Transfer
https://www.tensorflow.org/tutorials/generative/style_transfer

Neural Style Transfer

Source:TensorFlow Neural Style Transfer
https://www.tensorflow.org/tutorials/generative/style_transfer

Neural Style Transfer



Source:TensorFlow Neural Style Transfer
https://www.tensorflow.org/tutorials/generative/style_transfer

A Neural Algorithm of Artistic Style

Leon A. Gatys,[1,2,3*] Alexander S. Ecker,[1,2,4,5] Matthias Bethge[1,2,4]

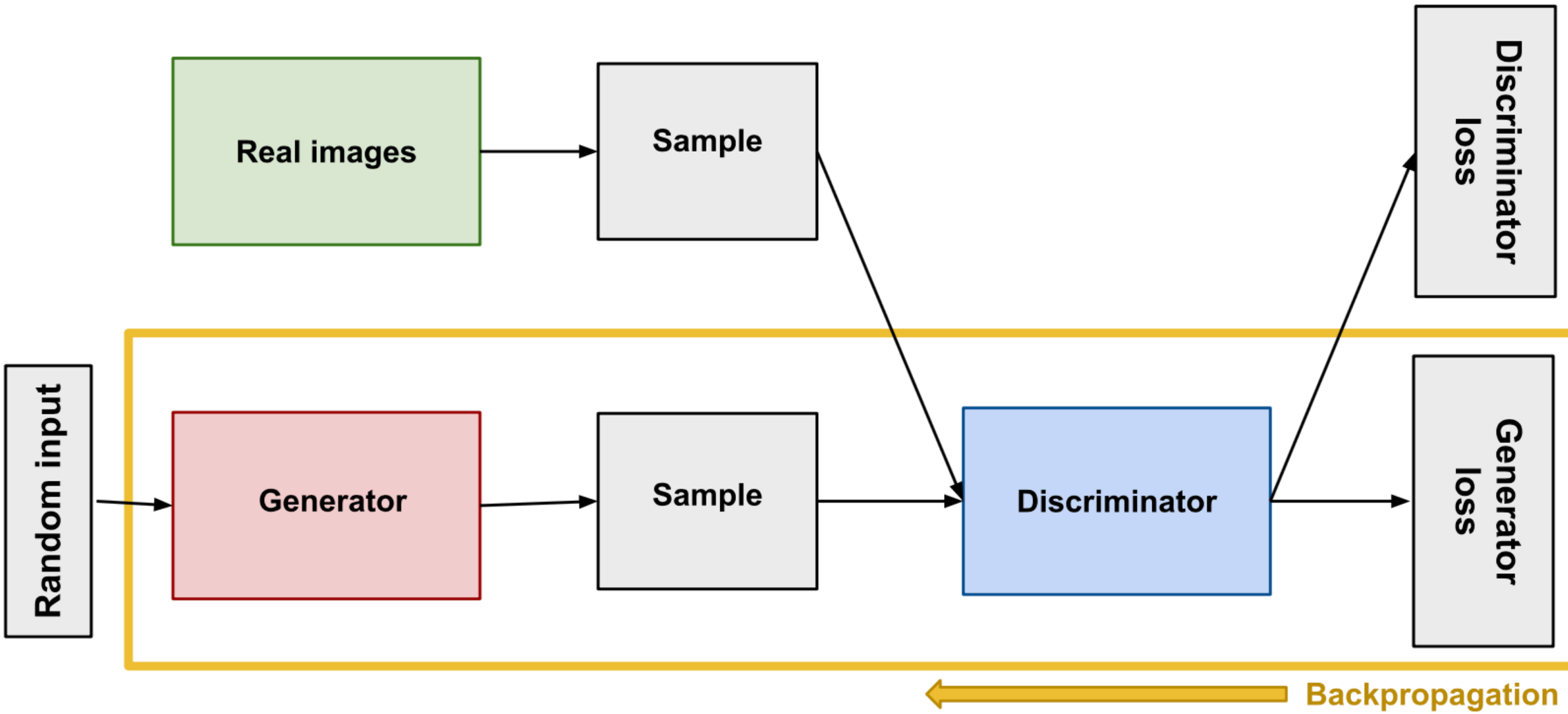
[1]Werner Reichardt Centre for Integrative Neuroscience
and Institute of Theoretical Physics, University of Tübingen, Germany
[2]Bernstein Center for Computational Neuroscience, Tübingen, Germany
[3]Graduate School for Neural Information Processing, Tübingen, Germany
[4]Max Planck Institute for Biological Cybernetics, Tübingen, Germany
[5]Department of Neuroscience, Baylor College of Medicine, Houston, TX, USA
*To whom correspondence should be addressed; E-mail: leon.gatys@bethgelab.org

In fine art, especially painting, humans have mastered the skill to create unique visual experiences through composing a complex interplay between the content and style of an image. Thus far the algorithmic basis of this process is unknown and there exists no artificial system with similar capabilities. However, in other key areas of visual perception such as object and face recognition near-human performance was recently demonstrated by a class of biologically inspired vision models called Deep Neural Networks.[1, 2] Here we introduce an artificial system based on a Deep Neural Network that creates artistic images

Generative Adversarial Networks (GANs)

- A Generative Adversarial Network (GAN) is a class of machine learning techniques in which two neural networks play against each other.
- The *generative network* generates candidates, while the *discriminative network* evaluate them.
- The generative network tries to create new samples that look similar from the true data (an original distribution, for example portraits). The goal of the discriminator is to identify if the data given in input are from the original distribution or not.
- The generative network's training objective is to increase the error rate of the discriminative network (i.e., to fool the discriminative network)
- Indeed, the discriminative network's training objective is to minimise its error rate in discriminating the input.
- This is used for images, videogame generation, scientific images, etc.

Generative Adversarial Networks (GANs)



Source: https://developers.google.com/machine-learning/gan/generator

Generative Adversarial Nets

Ian J. Goodfellow[*], **Jean Pouget-Abadie**[†], **Mehdi Mirza**, **Bing Xu**, **David Warde-Farley**, **Sherjil Ozair**[‡], **Aaron Courville**, **Yoshua Bengio**[§]
Département d'informatique et de recherche opérationnelle
Université de Montréal
Montréal, QC H3C 3J7

Abstract

We propose a new framework for estimating generative models via an adversarial process, in which we simultaneously train two models: a generative model G that captures the data distribution, and a discriminative model D that estimates the probability that a sample came from the training data rather than G. The training procedure for G is to maximize the probability of D making a mistake. This framework corresponds to a minimax two-player game. In the space of arbitrary functions G and D, a unique solution exists, with G recovering the training data distribution and D equal to $\frac{1}{2}$ everywhere. In the case where G and D are defined by multilayer perceptrons, the entire system can be trained with backpropagation. There is no need for any Markov chains or unrolled approximate inference networks during either training or generation of samples. Experiments demonstrate the potential of the framework through qualitative and quantitative evaluation of the generated samples.

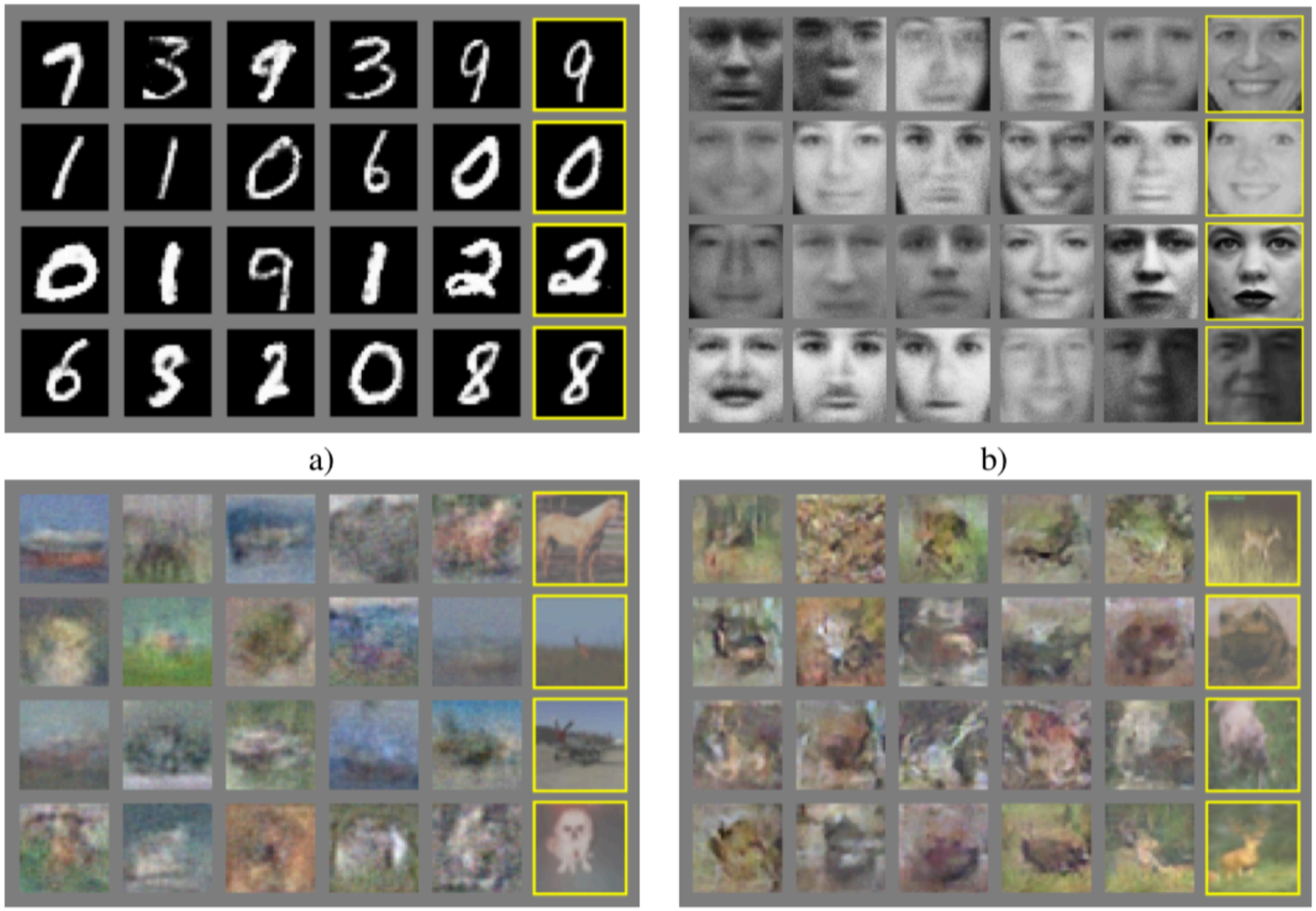
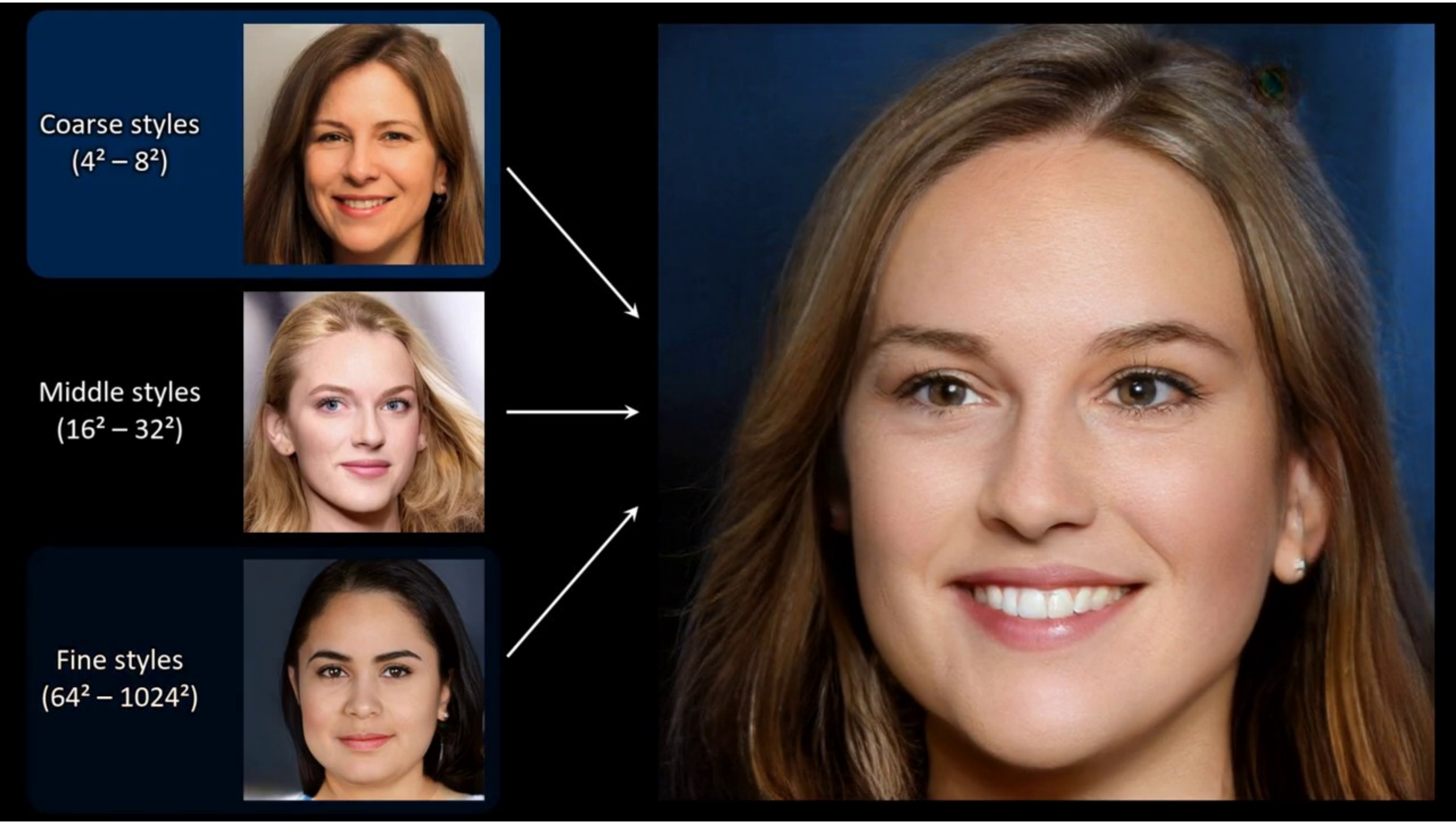


Figure 2: Visualization of samples from the model. Rightmost column shows the nearest training example of the neighboring sample, in order to demonstrate that the model has not memorized the training set. Samples are fair random draws, not cherry-picked. Unlike most other visualizations of deep generative models, these images show actual samples from the model distributions, not conditional means given samples of hidden units. Moreover, these samples are uncorrelated because the sampling process does not depend on Markov chain mixing. a) MNIST b) TFD c) CIFAR-10 (fully connected model) d) CIFAR-10 (convolutional discriminator and "deconvolutional" generator)

Generation of Images using Generative Adversarial Networks (GANs)

Source: David Foster. Generative Deep Learning. O'Reilly. 2019.

StyleGAN


Coarse styles
(4² – 8²)
Middle styles
(16² – 32²)
Fine styles
(64² – 1024²)

A Style-Based Generator Architecture for Generative Adversarial Networks

Tero Karras
NVIDIA
tkarras@nvidia.com

Samuli Laine
NVIDIA
slaine@nvidia.com

Timo Aila
NVIDIA
taila@nvidia.com

Abstract

We propose an alternative generator architecture for generative adversarial networks, borrowing from style transfer literature. The new architecture leads to an automatically learned, unsupervised separation of high-level attributes (e.g., pose and identity when trained on human faces) and stochastic variation in the generated images (e.g., freckles, hair), and it enables intuitive, scale-specific control of the synthesis. The new generator improves the state-of-the-art in terms of traditional distribution quality metrics, leads to demonstrably better interpolation properties, and also better disentangles the latent factors of variation. To quantify interpolation quality and disentanglement, we propose two new, automated methods that are applicable to any generator architecture. Finally, we introduce a new, highly varied and high-quality dataset of human faces.

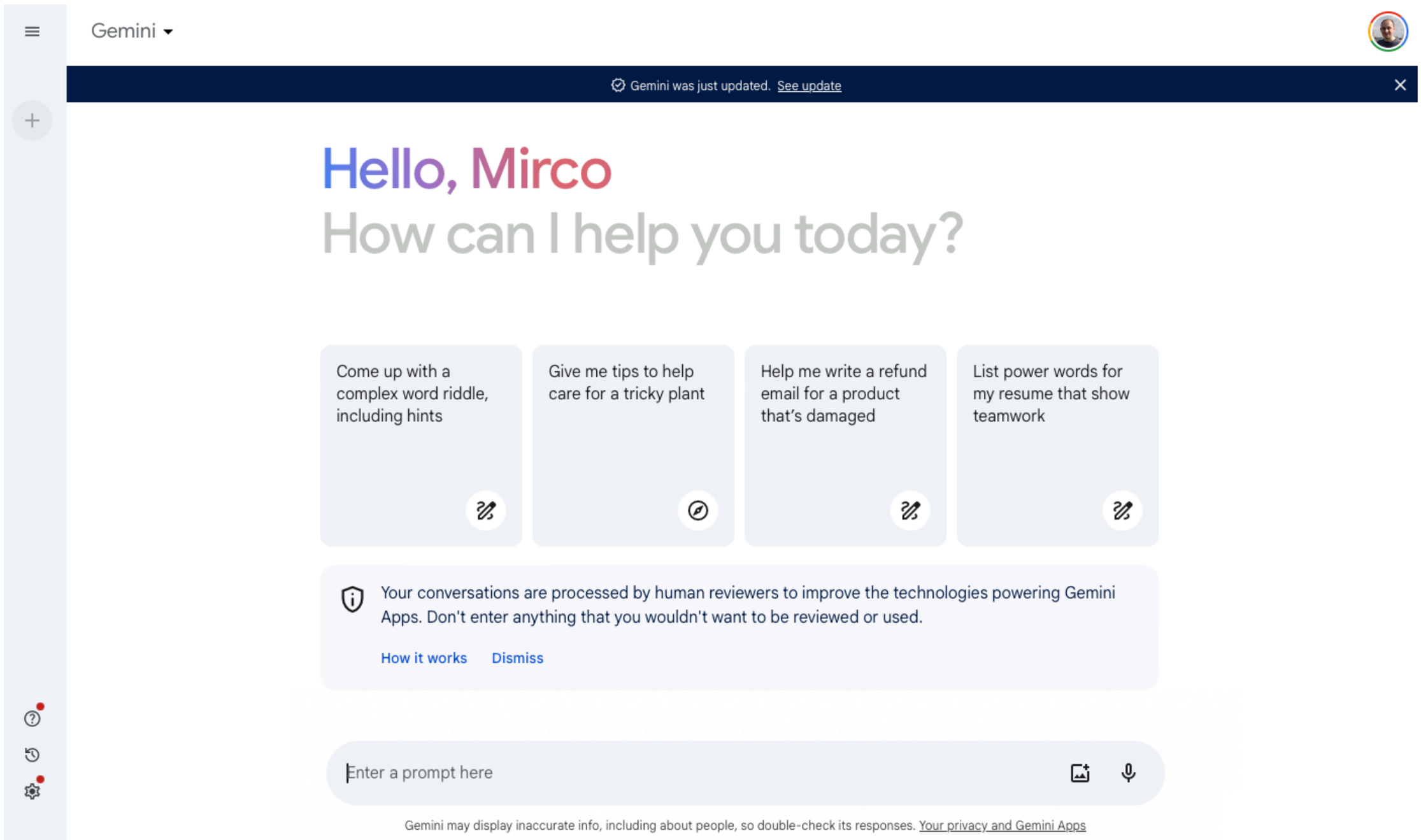
(e.g., pose, identity) from stochastic variation (e.g., freckles, hair) in the generated images, and enables intuitive scale-specific mixing and interpolation operations. We do not modify the discriminator or the loss function in any way, and our work is thus orthogonal to the ongoing discussion about GAN loss functions, regularization, and hyperparameters [24, 45, 5, 40, 44, 36].

Our generator embeds the input latent code into an intermediate latent space, which has a profound effect on how the factors of variation are represented in the network. The input latent space must follow the probability density of the training data, and we argue that this leads to some degree of unavoidable entanglement. Our intermediate latent space is free from that restriction and is therefore allowed to be disentangled. As previous methods for estimating the degree of latent space disentanglement are not directly applicable in our case, we propose two new automated metrics — perceptual path length and linear separability — for quantifying these aspects of the generator. Using these metrics, we show that compared to a traditional generator architecture,

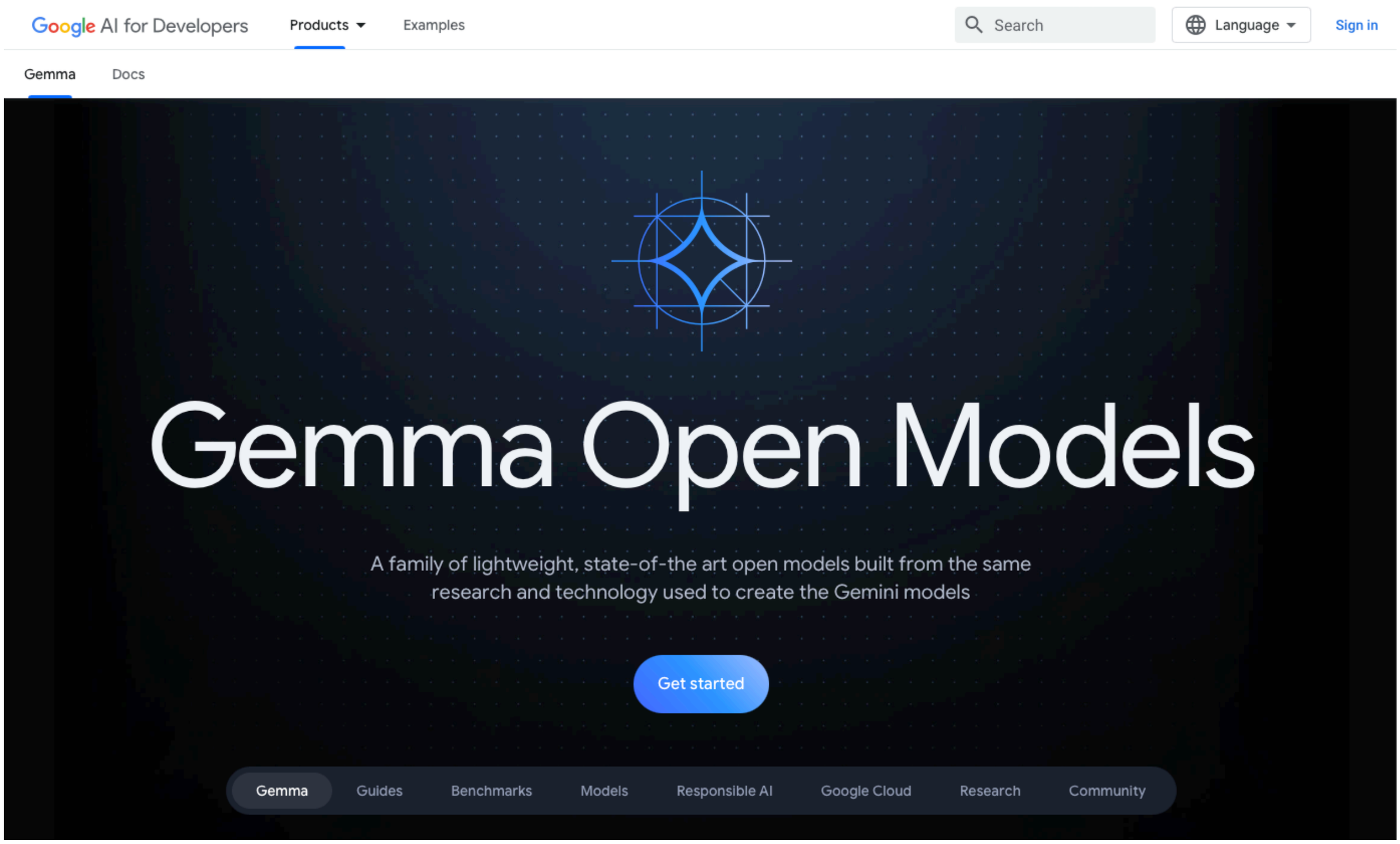
GPT-3 and GPT-4

- GPT-3 and GPT-4 (Generative Pre-Trained Transformer 3 and 4) language models that use deep learning to produce human-like text.
- Created by Open AI.
- They are based on the so-called *transformer architecture.*
- The GPT-3 model is based on 175 billion parameters.
- The GPT-4 model is based on 1 trillion parameters (rumours…).

Google Gemini



Google Gemma



Transformers

Attention Is All You Need

Ashish Vaswani* Google Brain avaswani@google.com

Noam Shazeer* Google Brain noam@google.com

Niki Parmar* Google Research nikip@google.com

Jakob Uszkoreit* Google Research usz@google.com

Llion Jones* Google Research llion@google.com

Aidan N. Gomez* † University of Toronto aidan@cs.toronto.edu

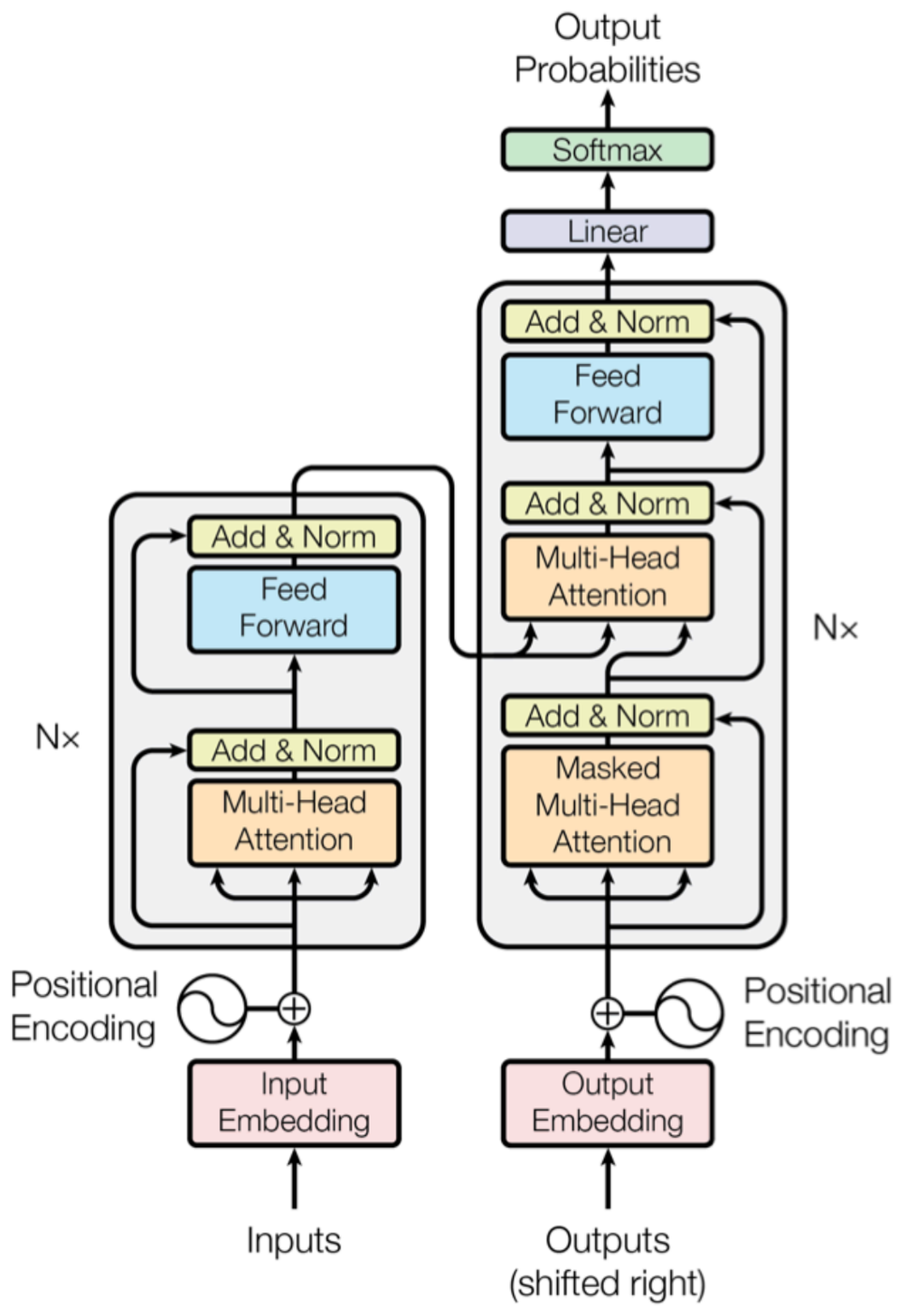
Łukasz Kaiser* Google Brain lukaszkaiser@google.com

Illia Polosukhin* ‡ illia.polosukhin@gmail.com

Abstract

The dominant sequence transduction models are based on complex recurrent or convolutional neural networks that include an encoder and a decoder. The best performing models also connect the encoder and decoder through an attention mechanism. We propose a new simple network architecture, the Transformer, based solely on attention mechanisms, dispensing with recurrence and convolutions entirely. Experiments on two machine translation tasks show these models to be superior in quality while being more parallelizable and requiring significantly less time to train. Our model achieves 28.4 BLEU on the WMT 2014 English-to-German translation task, improving over the existing best results, including ensembles, by over 2 BLEU. On the WMT 2014 English-to-French translation task, our model establishes a new single-model state-of-the-art BLEU score of 41.8 after training for 3.5 days on eight GPUs, a small fraction of the training costs of the best models from the literature. We show that the Transformer generalizes well to other tasks by applying it successfully to English constituency parsing both with large and limited training data.

Transformers



DALL·E

- Introduced in January 2021.
- DALL E is a 12-billion parameter version of GPT-3 trained to generate images from text descriptions using a dataset of text-image pairs.
- Different capabilities including:
  - Creation of anthropomorphised versions of animals and objects;
  - Linking unrelated concepts in novel ways;
  - Text rendering;
  - Transformation of existing images;
  - …

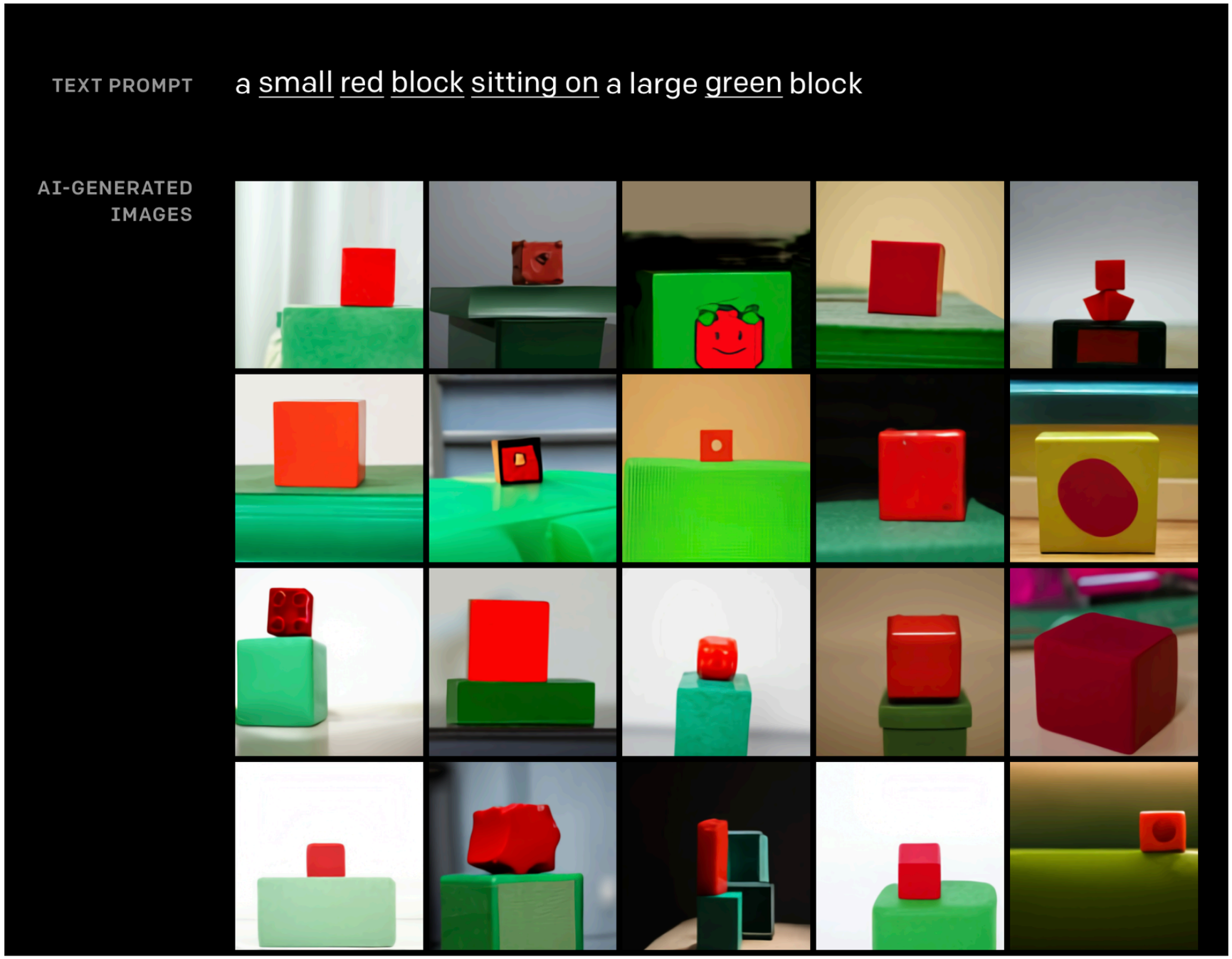
DALL·E

TEXT PROMPT

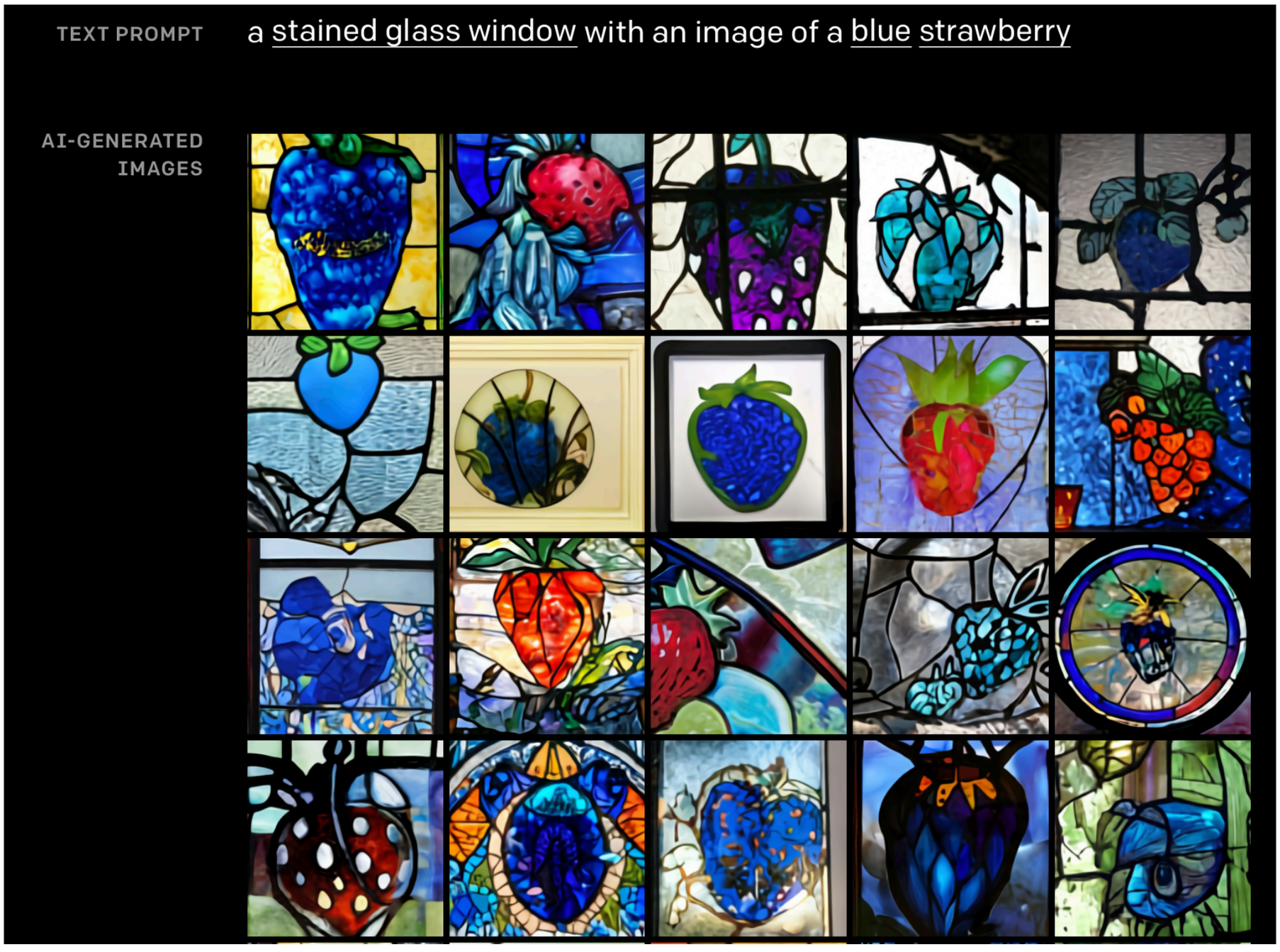
an armchair in the shape of an avocado. an armchair imitating an avocado.

AI-GENERATED IMAGES

DALL·E


TEXT PROMPT
a small red block sitting on a large green block
AI-GENERATED IMAGES

DALL·E


TEXT PROMPT
a stained glass window with an image of a blue strawberry
AI-GENERATED IMAGES

DALL·E

Zero-Shot Text-to-Image Generation

Aditya Ramesh [1] **Mikhail Pavlov** [1] **Gabriel Goh** [1] **Scott Gray** [1]
Chelsea Voss [1] **Alec Radford** [1] **Mark Chen** [1] **Ilya Sutskever** [1]

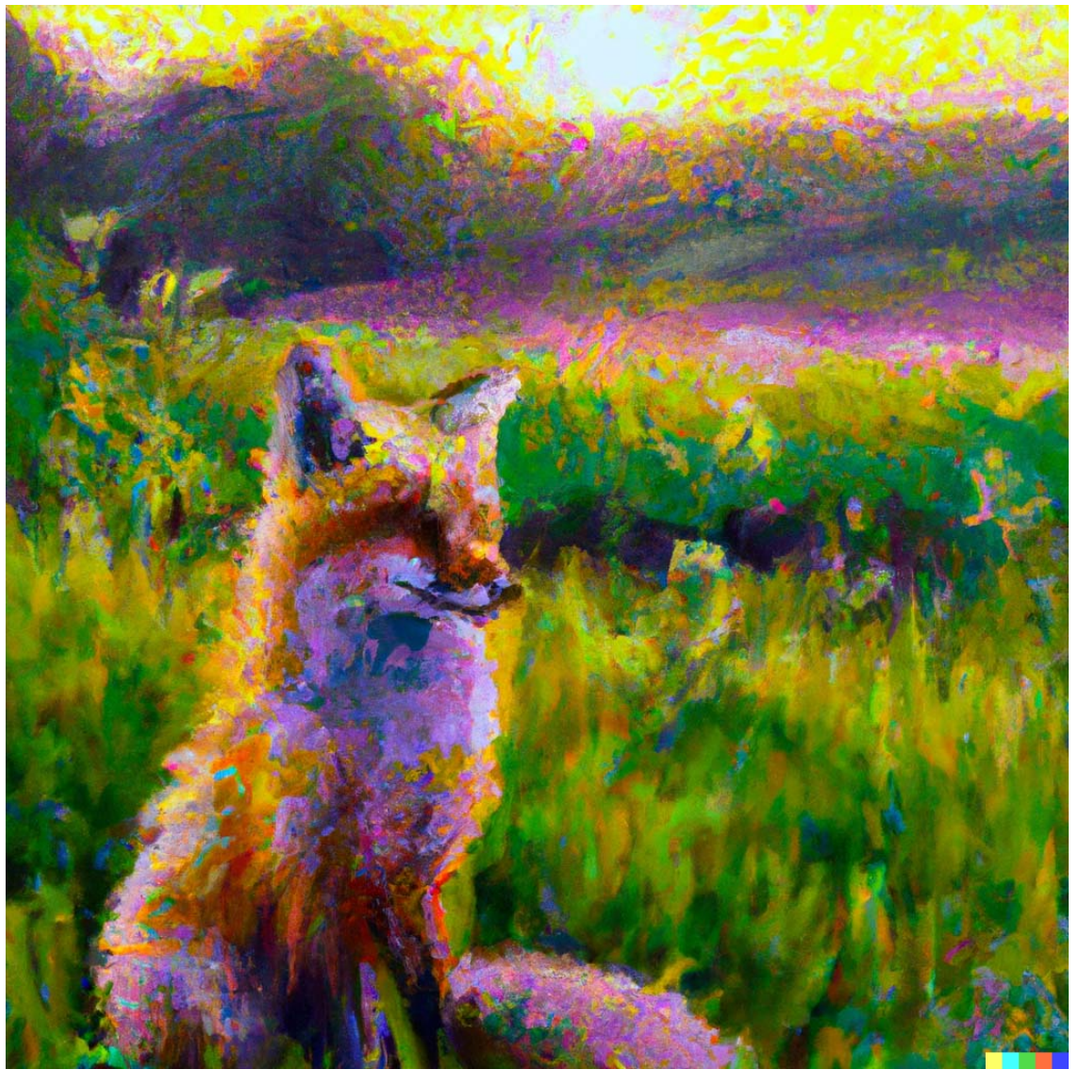
Abstract

Text-to-image generation has traditionally focused on finding better modeling assumptions for training on a fixed dataset. These assumptions might involve complex architectures, auxiliary losses, or side information such as object part labels or segmentation masks supplied during training. We describe a simple approach for this task based on a transformer that autoregressively models the text and image tokens as a single stream of data. With sufficient data and scale, our approach is competitive with previous domain-specific models when evaluated in a zero-shot fashion.

Figure 1. Comparison of original images (top) and reconstructions

DALL·E 2

DALL·E 1



DALL·E 2

“a painting of a fox sitting in a field at sunrise in the style of Claude Monet”

DALL·E 2

- Launched in April 2022 - Dall·E is able to create “original, realistic images and art from a text description.”
- It is able to combine concept attributes and styles with visible improvements with respect to the previous version of the system.
- It can be used for photorealistic editing and creation of variations with high resolution.
- It learns the relationships between the images and the text used to describe them, including abstract ones.

DALL·E 2

"an astronaut riding a horse lounging in a tropical resort in space in a photorealistic style”

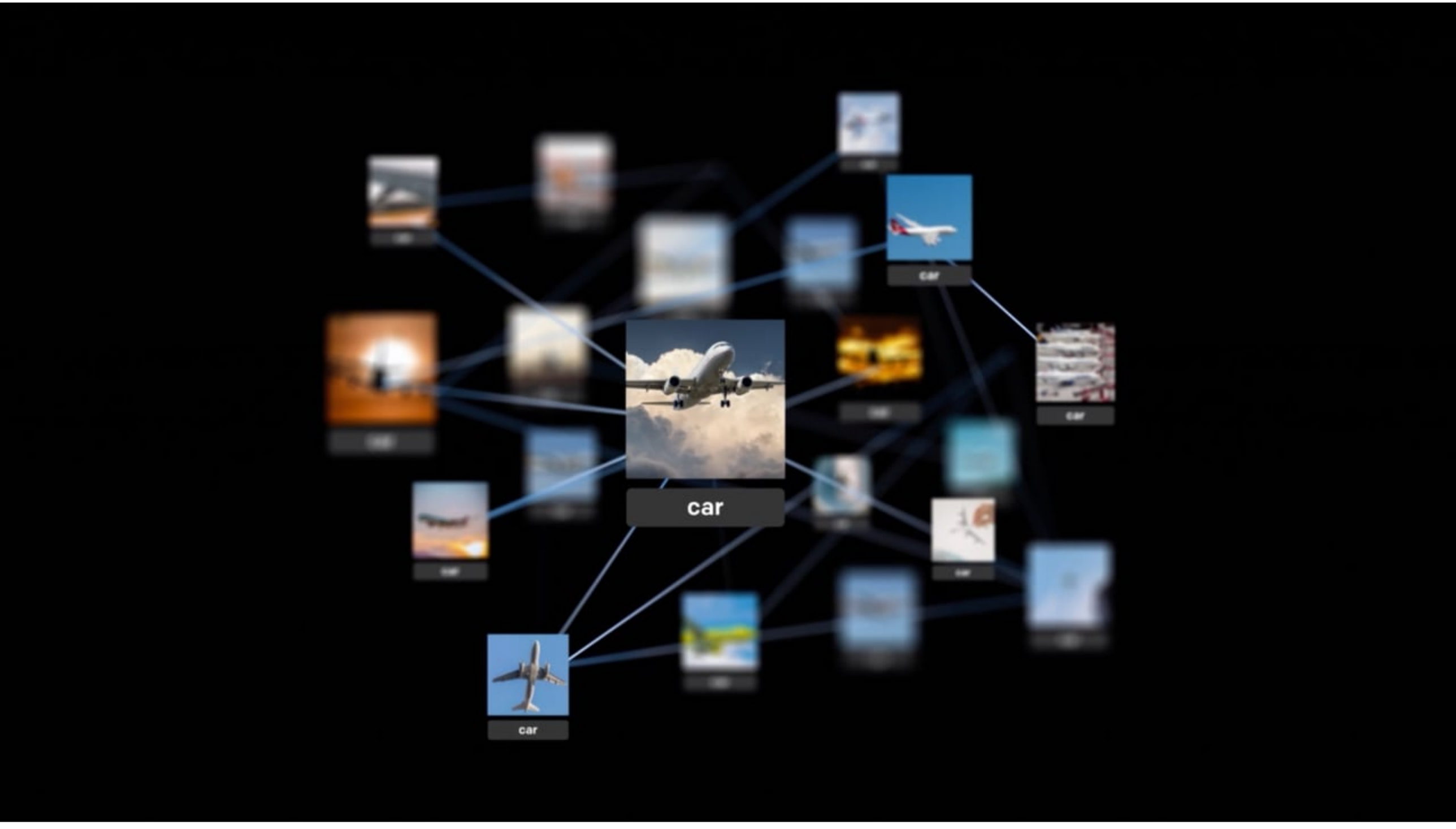
DALL·E 2

DALL·E 2






Variations of Vermeer’s “Girl with a Pearl Earring”


car
car
car
car
car

DALL·E 2

Hierarchical Text-Conditional Image Generation with CLIP Latents

Aditya Ramesh* OpenAI aramesh@openai.com

Prafulla Dhariwal* OpenAI prafulla@openai.com

Alex Nichol* OpenAI alex@openai.com

Casey Chu* OpenAI casey@openai.com

Mark Chen OpenAI mark@openai.com

Abstract

Contrastive models like CLIP have been shown to learn robust representations of images that capture both semantics and style. To leverage these representations for image generation, we propose a two-stage model: a prior that generates a CLIP image embedding given a text caption, and a decoder that generates an image conditioned on the image embedding. We show that explicitly generating image representations improves image diversity with minimal loss in photorealism and caption similarity. Our decoders conditioned on image representations can also produce variations of an image that preserve both its semantics and style, while varying the non-essential details absent from the image representation. Moreover, the joint embedding space of CLIP enables language-guided image manipulations in a zero-shot fashion. We use diffusion models for the decoder and experiment

DALL·E 3

“A folk music band composed of anthropomorphic autumn leaves, each playing traditional bluegrass instruments, amidst a rustic setting dappled with the soft light of a harvest moon. ”

DALL·E 3

Improving Image Generation with Better Captions

James Betker[*†] jbetker@openai.com
Gabriel Goh[*†] ggoh@openai.com
Li Jing[*†] lijing@openai.com
Tim Brooks[†]

Jianfeng Wang[‡] **Linjie Li**[‡] **Long Ouyang**[†] **Juntang Zhuang**[†] **Joyce Lee**[†] **Yufei Guo**[†]

Wesam Manassra[†] **Prafulla Dhariwal**[†] **Casey Chu**[†] **Yunxin Jiao**[†]

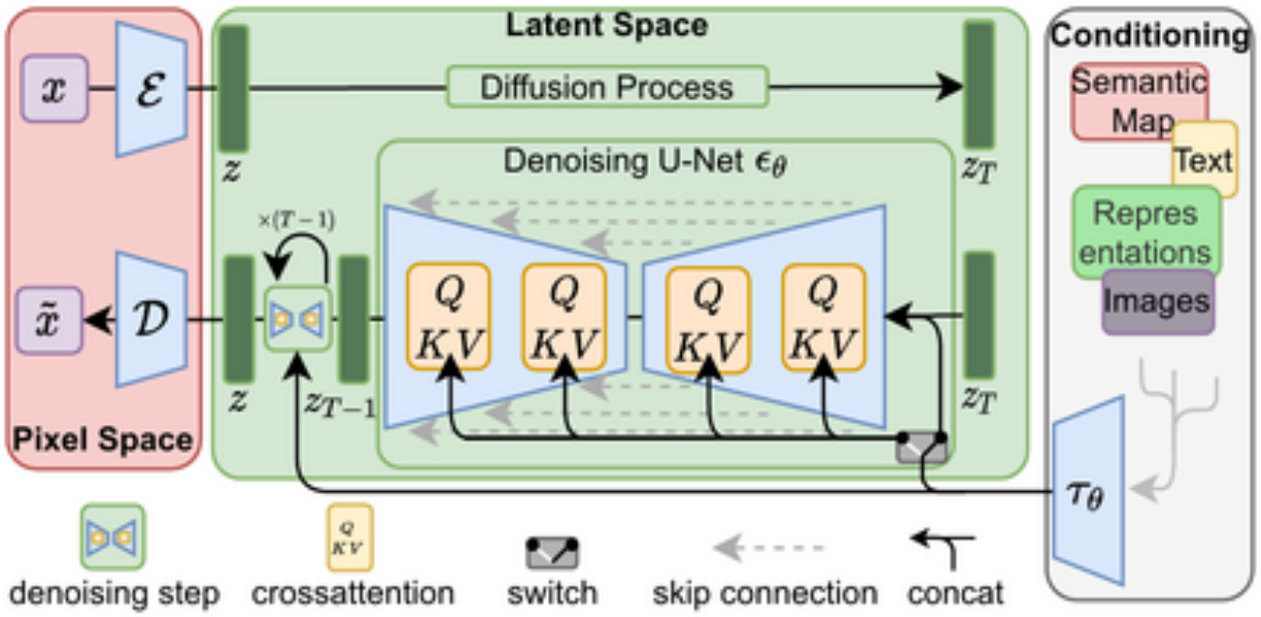
Aditya Ramesh[*†] aramesh@openai.com

Abstract

We show that prompt following abilities of text-to-image models can be substantially improved by training on highly descriptive generated image captions. Existing text-to-image models struggle to follow detailed image descriptions and often ignore words or confuse the meaning of prompts. We hypothesize that this issue stems from noisy and inaccurate image captions in the training dataset. We address this by training a bespoke image captioner and use it to recaption the training dataset. We then train several text-to-image models and find that training on these synthetic captions reliably improves prompt following ability. Finally, we use these findings to build DALL-E 3: a new text-to-image generation system, and benchmark its performance on an evaluation designed to measure prompt following, coherence, and aesthetics, finding that it compares favorably to competitors. We publish samples and code for these evaluations so that future research can continue optimizing this important aspect of text-to-image systems.

Stable Diffusion

Stable Diffusion



Source: Wikimedia

- It is based on an autoencoder architecture based on *denoising* modules conditioned on multimodal inputs.
- The code is publicly available.
- Initially developed by LMU, commercialised by Stability AI.

Deep Reinforcement Learning from Human Preferences

Paul F Christiano
OpenAI
paul@openai.com

Jan Leike
DeepMind
leike@google.com

Tom B Brown
Google Brain*
tombbrown@google.com

Miljan Martic
DeepMind
miljanm@google.com

Shane Legg
DeepMind
legg@google.com

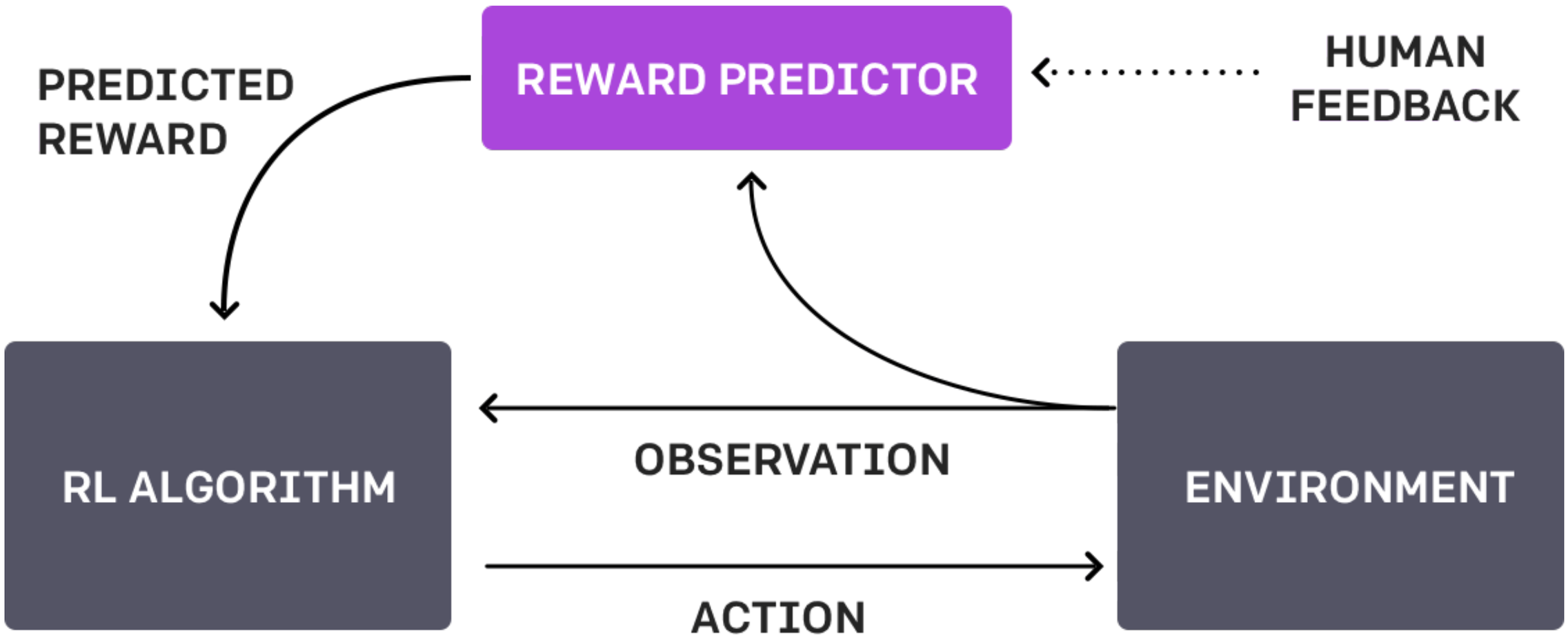
Dario Amodei
OpenAI
damodei@openai.com

Abstract

For sophisticated reinforcement learning (RL) systems to interact usefully with real-world environments, we need to communicate complex goals to these systems. In this work, we explore goals defined in terms of (non-expert) human preferences between pairs of trajectory segments. We show that this approach can effectively solve complex RL tasks without access to the reward function, including Atari games and simulated robot locomotion, while providing feedback on less than 1% of our agent's interactions with the environment. This reduces the cost of human oversight far enough that it can be practically applied to state-of-the-art RL systems. To demonstrate the flexibility of our approach, we show that we can successfully train complex novel behaviors with about an hour of human time. These behaviors and environments are considerably more complex than any which have been previously learned from human feedback.

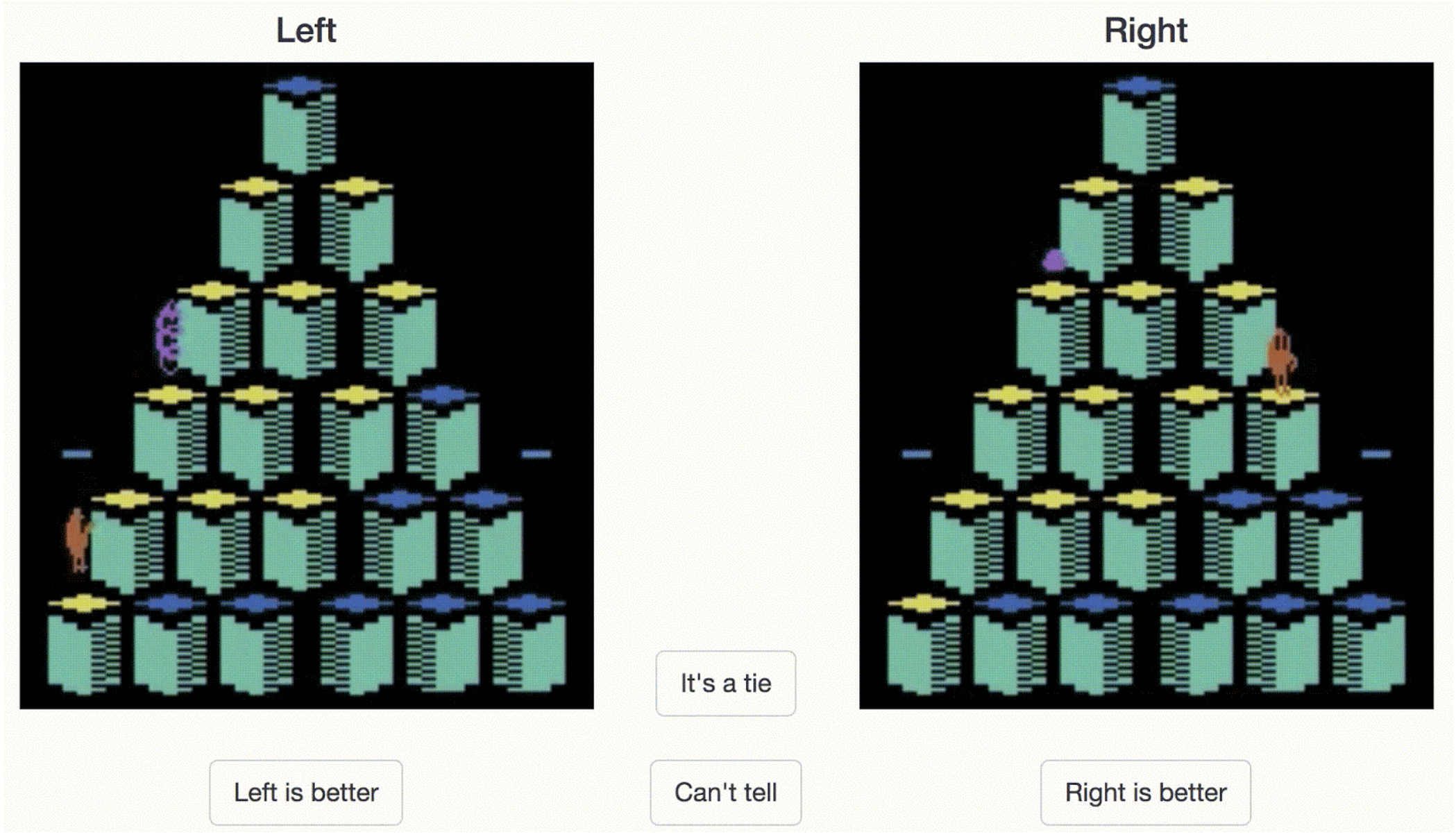
NeurIPS 2017

Reinforcement Learning from Human Feedback (RLHF)



Source: OpenAI

Reinforcement Learning from Human Feedback (RLHF)



Source: DeepMind

Generative Agents

Generative Agents: Interactive Simulacra of Human Behavior

Joon Sung Park
Stanford University
Stanford, USA
joonspk@stanford.edu

Joseph C. O'Brien
Stanford University
Stanford, USA
jobrien3@stanford.edu

Carrie J. Cai
Google Research
Mountain View, CA, USA
cjcai@google.com

Meredith Ringel Morris
Google DeepMind
Seattle, WA, USA
merrie@google.com

Percy Liang
Stanford University
Stanford, USA
pliang@cs.stanford.edu

Michael S. Bernstein
Stanford University
Stanford, USA
msb@cs.stanford.edu

Figure 1: Generative agents are believable simulacra of human behavior for interactive applications. In this work, we demonstrate generative agents by populating a sandbox environment, reminiscent of The Sims, with twenty-five agents. Users can observe and intervene as agents plan their days, share news, form relationships, and coordinate group activities.

ACM UIST 2023

What is Creativity?

Source: Wikimedia

“Creativity can be defined as the ability to generate novel, and valuable, ideas. Valuable, here, has many meanings: interesting, useful, beautiful, simple, richly complex, and so on. Ideas covers many meaning too: not only ideas as such (concepts, theories, interpretations, stories), but also artifacts such as graphic images, sculptures, houses and jet engines. Computer models have been designed to generate ideas in all these areas and more.”

Margaret A. Boden

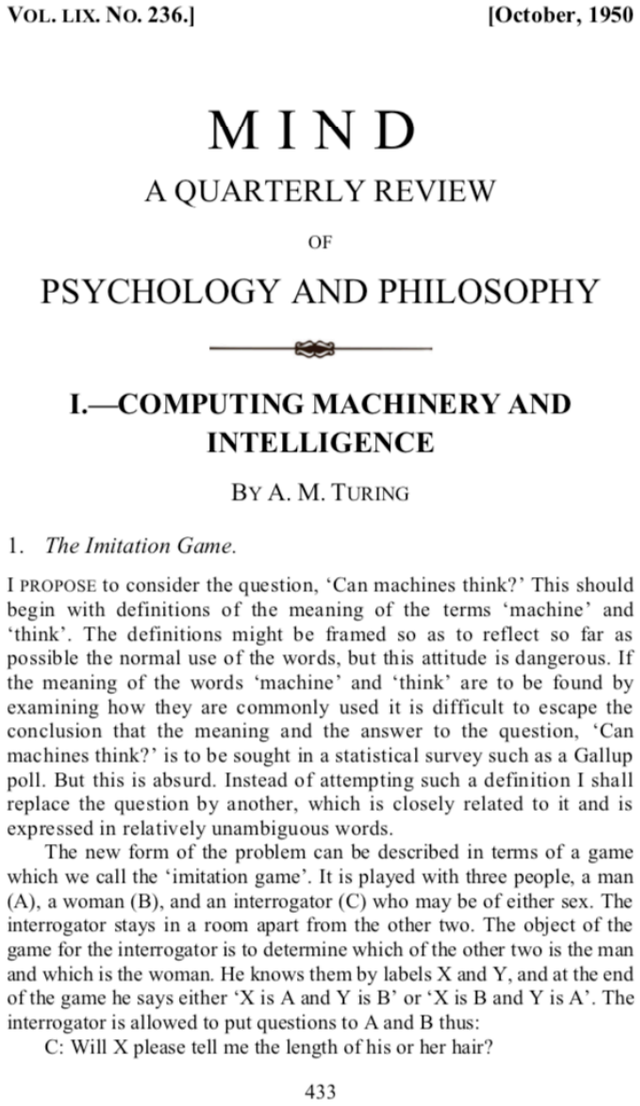
Ada Lovelace's Objection

Source: Computer History Museum

"The Analytical Engine has no pretensions whatever to originate anything. It can do whatever we know how to order it to perform; but it has no power of anticipating any analytical relations or truths."

Ada Lovelace

Turing's Response



VOL. LIX. NO. 236.] [October, 1950

MIND

A QUARTERLY REVIEW

OF

PSYCHOLOGY AND PHILOSOPHY

I.—COMPUTING MACHINERY AND INTELLIGENCE

BY A. M. TURING

1. *The Imitation Game.*

I PROPOSE to consider the question, 'Can machines think?' This should begin with definitions of the meaning of the terms 'machine' and 'think'. The definitions might be framed so as to reflect so far as possible the normal use of the words, but this attitude is dangerous. If the meaning of the words 'machine' and 'think' are to be found by examining how they are commonly used it is difficult to escape the conclusion that the meaning and the answer to the question, 'Can machines think?' is to be sought in a statistical survey such as a Gallup poll. But this is absurd. Instead of attempting such a definition I shall replace the question by another, which is closely related to it and is expressed in relatively unambiguous words.

The new form of the problem can be described in terms of a game which we call the 'imitation game'. It is played with three people, a man (A), a woman (B), and an interrogator (C) who may be of either sex. The interrogator stays in a room apart from the other two. The object of the game for the interrogator is to determine which of the other two is the man and which is the woman. He knows them by labels X and Y, and at the end of the game he says either 'X is A and Y is B' or 'X is B and Y is A'. The interrogator is allowed to put questions to A and B thus:

C: Will X please tell me the length of his or her hair?

433

- Ada Lovelace's objection can be seen as the assertion that computers cannot surprise us.
- Alan Turing in his Mind paper argues that actually computers are still able to surprise us. He also underlines the fact that Ada Lovelace lived in a period where neurological phenomena were not known.

Can an (artificial) agent be creative?

References

- Francois Chollet. Deep Learning with Python. Second Edition. Manning. 2022.

- David Foster. Generative Deep Learning. Second Edition. O'Reilly. 2023.

- Arthur I. Miller. The Artist in the Machine. The World of AI-Powered Creativity. MIT Press. 2019.