



GEN AI?

What is the influence of AI on the young generation of scientists?

Cyrus Walther [WG16, CODATA, TU Dortmund University]

IUPAP Inter-Commission Symposium

(THE) ARTIFICIAL INTELLIGENCE, [NOUN]

(THE) АРТИФИЦИАЛ ИНТЕЛЛИГЕНС' [ИОНИ]

- ◆ *The capacity of computers or other machines to exhibit or simulate intelligent behaviour.*
~ *Oxford English Dictionary*
- ◆ *AI is generally considered to be a discipline of computer science that is aimed at developing machines and systems that can carry out tasks considered to require human intelligence.*
~ *UN Regional Information Centre for Western Europe*

Artificial Intelligence - Buzzword Bingo?

M

Machine Learning

Enabling AI to imitate the way that humans learn

N

Neuromorphic Computing

Developing computing that mimics the concept in which the human brain thinks

N

Neural Networks

Making decisions by using processes that mimic the way biological neurons work together

D

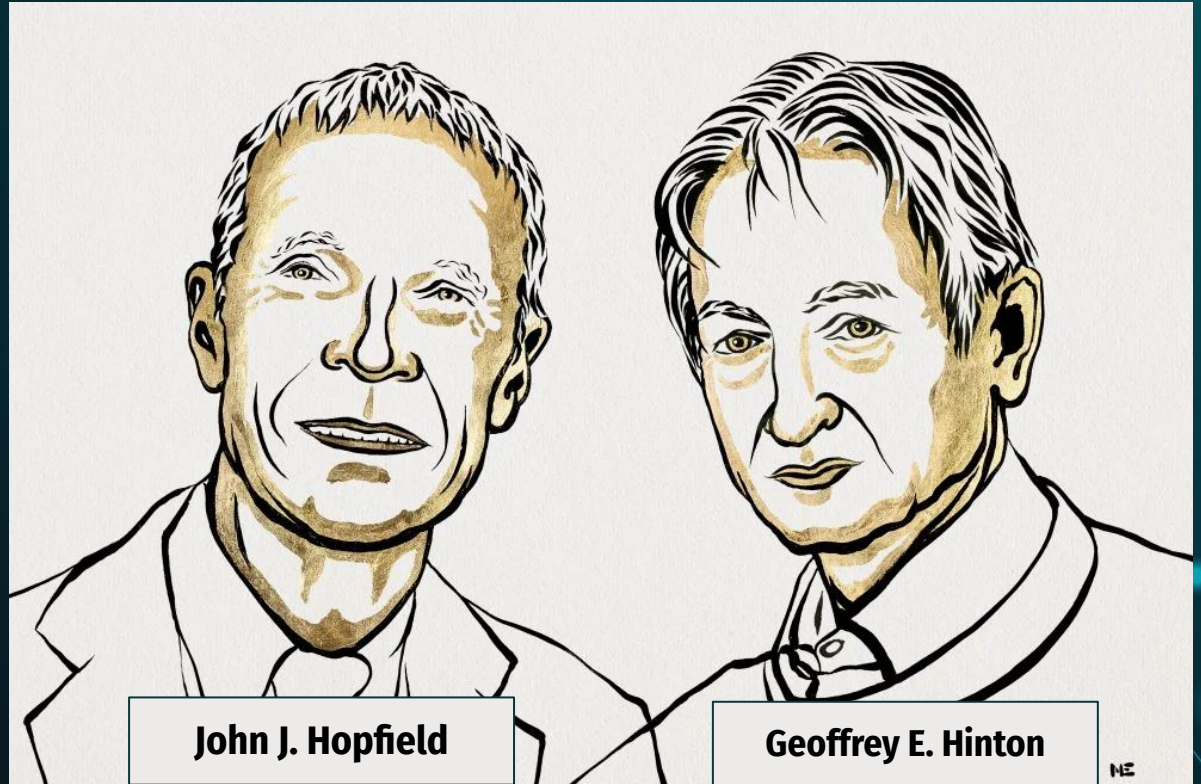
Deep Learning

Using multilayer Neural Networks to stimulate the complex decision-making of the human brain

And it is timely...

***The 2024 Nobel
Prize in Physics***

***"for foundational
discoveries and
inventions that
enable machine
learning with
artificial neural
networks"***



John J. Hopfield

Geoffrey E. Hinton

And it comes in with momentum...

High-Resolution Image Synthesis with Latent Diffusion Models

Robin Rombach^{1*}, Andreas Blattmann^{2*}, Dominik Lorenz¹, Patrick Esser³, Björn Ommer¹

¹Ludwig Maximilian University of Munich & IWR, Heidelberg University, Germany ²Roway ML

<https://github.com/CompVis/latent-diffusion>

Abstract

By decomposing the image formation process into a sequential application of denoising autoencoders, diffusion models (DMs) achieve state-of-the-art synthesis results on image data and beyond. Additionally, their formulation allows for a guiding mechanism to control the image generation process without retraining. However, since these models typically operate directly in pixel space, optimization of powerful DMs often consumes hundreds of GPU days and inference is expensive due to sequential evaluations. To enable DM training on limited computational resources while retaining their quality and flexibility, we apply them in the latent space of powerful pretrained autoencoders. In contrast to previous work, training diffusion models on such a representation allows for the first time to reach a near-optimal balance between complexity reduction and detail preservation, greatly boosting visual fidelity. By introducing cross-attention layers into the model architecture, we turn diffusion models into powerful and flexible generators for general conditioning inputs such as text or bounding boxes and high-resolution synthesis becomes possible in a convolutional manner. Our latent diffusion models (LDMs) achieve new state-of-the-art scores for image inpainting and class-conditional image synthesis and highly competitive performance on various tasks, including text-to-image synthesis, unconditional image generation and super-resolution, while significantly reducing computational requirements compared to pixel-based DMs.

1. Introduction

Image synthesis is one of the computer vision fields with the most spectacular recent development, but also among those with the greatest computational demands. Especially high-resolution synthesis of complex, natural scenes is presently dominated by scaling up likelihood-based models, potentially containing billions of parameters in autoregressive (AR) transformers [66, 67]. In contrast, the promising results of GANs [2, 27, 40] have been revealed to be mostly confined to data with comparably limited variability as their adversarial learning procedure does not easily scale to modeling complex, multi-modal distributions. Recently, diffusion models [82], which are built from a hierarchy of denoising autoencoders, have shown to achieve impressive

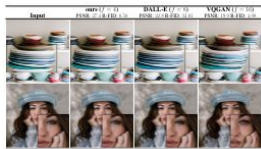


Figure 1. Boosting the upper bound on achievable quality with less aggressive downsampling. Since diffusion models offer excellent inductive biases for spatial data, we do not need the heavy spatial downsampling of related generative models in latent space, but can still greatly reduce the dimensionality of the data via suitable autoencoding models, see Sec. 3. Images are from the DIV2K [1] validation set, evaluated at 512^2 px. We denote the spatial downsampling factor by f . Reconstruction FID₁₀₀ and PSNR are calculated on ImageNet-val. [12]; see also Tab. 8.

results in image synthesis [30,85] and beyond [7,45,48,57], and define the state-of-the-art in class-conditional image synthesis [15,31] and super-resolution [72]. Moreover, even unconditional DMs can readily be applied to tasks such as inpainting and colorization [85] or stroke-based synthesis [33], in contrast to other types of generative models [19,46,69]. Being likelihood-based models, they do not exhibit mode-collapse and training instabilities as GANs and, by heavily exploiting parameter sharing, they can model highly complex distributions of natural images without involving billions of parameters as in AR models [67].

Democratizing High-Resolution Image Synthesis. DMs belong to the class of likelihood-based models, whose mode-covering behavior makes them prone to spend excessive amounts of capacity (and thus compute resources) on modeling imperceptible details of the data [16, 73]. Although the reweighted variational objective [30] aims to address this by undersampling the initial denoising steps, DMs are still computationally demanding, since training and evaluating such a model requires repeated function evaluations (and gradient computations) in the high-dimensional space of RGB images. As an example, training the most powerful DMs often takes hundreds of GPU days (e.g. 150–1000 V100 days in [51]) and repeated evaluations on a noisy version of the input space render also inference expensive,

izing Flows

DANILOR@GOOGLE.COM
SHAKIR@GOOGLE.COM

and chemical systems. Despite these surprising advances, there are a number of disadventurous methods that limit their power and wider adoption as a default method for statistical inference. One of these limitations, the choice of approximation, that we address in this paper.

DMs require that intractable posterior distributions be approximated by a class of known probability distributions τ which we search for the best approximation. The class of approximations used, e.g., mean-field approximations, implies that inference is able to resemble the true posterior is a widely raised objection to variational inference and alternative methods such as the asymptotic regime we are unable to retrain.

DMs require that intractable posterior distributions be approximated by a class of known probability distributions τ which we search for the best approximation. The class of approximations used, e.g., mean-field approximations, implies that inference is able to resemble the true posterior is a widely raised objection to variational inference and alternative methods such as the asymptotic regime we are unable to retrain.

DMs require that intractable posterior distributions be approximated by a class of known probability distributions τ which we search for the best approximation. The class of approximations used, e.g., mean-field approximations, implies that inference is able to resemble the true posterior is a widely raised objection to variational inference and alternative methods such as the asymptotic regime we are unable to retrain.

Graph Domains

Franco Scarselli
Dipartimento di Informatica
Università di Siena, Italy
E-mail: franco@di.unisi.it

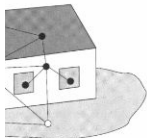


Figure 2. Graphical representation by a RAG.

Graph domains (RNNs [31, 44] are a new neural network this problem. In fact, RNNs can be used to encode the data into a set of states associated with the nodes, which are dynamically updated following the edges among the nodes. Finally, an output is generated based on the states stored in the nodes. However, RNNs have a number of limitations. In fact, they are directed and acyclic graphs and can focus on problems, i.e. $\tau(G, n)$ must be

DMs require that intractable posterior distributions be approximated by a class of known probability distributions τ which we search for the best approximation. The class of approximations used, e.g., mean-field approximations, implies that inference is able to resemble the true posterior is a widely raised objection to variational inference and alternative methods such as the asymptotic regime we are unable to retrain.

Graph Neural Networks

Graph neural networks (GNNs) represent the module or the cardinality of the graph. It is applied on a real number. The norm one of vector e is denoted

ial Nets

Bing Xu, David Warde-Farley, Shua Bengio
he operational

DMs require that intractable posterior distributions be approximated by a class of known probability distributions τ which we search for the best approximation. The class of approximations used, e.g., mean-field approximations, implies that inference is able to resemble the true posterior is a widely raised objection to variational inference and alternative methods such as the asymptotic regime we are unable to retrain.

DMs require that intractable posterior distributions be approximated by a class of known probability distributions τ which we search for the best approximation. The class of approximations used, e.g., mean-field approximations, implies that inference is able to resemble the true posterior is a widely raised objection to variational inference and alternative methods such as the asymptotic regime we are unable to retrain.

DMs require that intractable posterior distributions be approximated by a class of known probability distributions τ which we search for the best approximation. The class of approximations used, e.g., mean-field approximations, implies that inference is able to resemble the true posterior is a widely raised objection to variational inference and alternative methods such as the asymptotic regime we are unable to retrain.

DMs require that intractable posterior distributions be approximated by a class of known probability distributions τ which we search for the best approximation. The class of approximations used, e.g., mean-field approximations, implies that inference is able to resemble the true posterior is a widely raised objection to variational inference and alternative methods such as the asymptotic regime we are unable to retrain.

DMs require that intractable posterior distributions be approximated by a class of known probability distributions τ which we search for the best approximation. The class of approximations used, e.g., mean-field approximations, implies that inference is able to resemble the true posterior is a widely raised objection to variational inference and alternative methods such as the asymptotic regime we are unable to retrain.

Models

Pieter Abbeel
UC Berkeley
e18@cs.berkeley.edu

DMs require that intractable posterior distributions be approximated by a class of known probability distributions τ which we search for the best approximation. The class of approximations used, e.g., mean-field approximations, implies that inference is able to resemble the true posterior is a widely raised objection to variational inference and alternative methods such as the asymptotic regime we are unable to retrain.

DMs require that intractable posterior distributions be approximated by a class of known probability distributions τ which we search for the best approximation. The class of approximations used, e.g., mean-field approximations, implies that inference is able to resemble the true posterior is a widely raised objection to variational inference and alternative methods such as the asymptotic regime we are unable to retrain.



Figure 3. Samples from a wide variety of generative models, including text-to-image, audio-to-image, and audio-to-audio.

DMs require that intractable posterior distributions be approximated by a class of known probability distributions τ which we search for the best approximation. The class of approximations used, e.g., mean-field approximations, implies that inference is able to resemble the true posterior is a widely raised objection to variational inference and alternative methods such as the asymptotic regime we are unable to retrain.

*The first two authors contributed equally to this work.

A GLANCE AT THE GEN AI TIMELINE

Snapchat



2015

Snapchat launched
AI-powered Filters

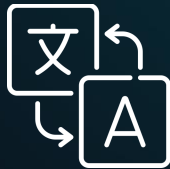


DeepL



08/2017

Translational AI DeepL
was released



Github Co-Pilot



11/2021

Release of the coding
AI Github Copilot



ChatGPT 3.0



11/2022

Release of the
conversational AI
ChatGPT 3.0



**Why
shouldn't we...**

**Write proposals
through AI?**

**Why
shouldn't we...**

**Stop handmade code
development?**

**Why
shouldn't we...**

**Create AI conference
summaries and stop
conference attendance?**

**So why
shouldn't we?**



Arguments that encourage the Use of AI



The young generation adapts to new technology quickly and identifies its capabilities in the process



The use of AI appears natural to the young community



Peer pressure enforces a wide coverage of users



Little regulation opens the possibility to usage in a variety of ways



It might be the Holy Grail for a relaxed life!

Arguments that keep us from the Use of AI

Loss of individuality through the usage of AI and rise of self doubt of personal skills



Reputation of AI remains to be disputable



AI learns underlying concepts and develops prediction power but extracting these physical concepts is complex



Science is about struggling and problem solving



Consequently, we end up visibly unsure!

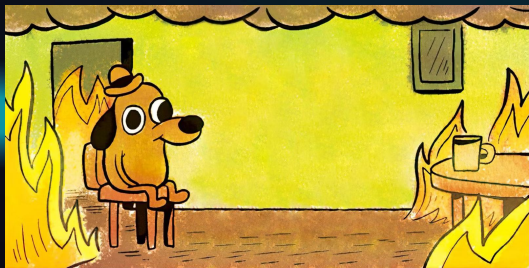


There are conflicts arising... - A Thought Process

I need to advance toward a permanent position as soon as possible!

Well, if I use AI, am I even good enough to climb the academic career path without it?

Well, if I have to rely on it to perform, then does it even rely on me?



AI tools will make my proposals better and make me receive this vital funding, so I have to use it!

But if everybody else uses it, I won't advance and be able to compete!

Am I even needed and is there any future for me in my passion?

What can we do? - We need a Strategic Approach!

"It's not about allowance or restriction but about guidance."



CONSCIOUSNESS



WELLBEING



EDUCATION



AWARENESS



Summarizing...

"It is our Young Generation that is heavily impacted by Artificial Intelligence. But caught between Peer Pressure and Loss of Identity, they lose navigation. Therefore, it is for all of us to become aware of this situation and educate ourselves on this subject. We have to develop an aura of wellbeing around the topic Artificial Intelligence and foster consciousness on the use of AI among students and among us!"

Thank you for your valuable attention!



I hope it was also valuable for you!