Stale Diffusion: Hyper-Realistic 5D Movie Generation using Old-school Methods

João F. Henriques,[*] **Dylan Campbell,**[†] **Tengda Han,**[‡] [§]
International Centre for Sleep Studies
Nappington, UK

Abstract

Two years ago, Stable Diffusion achieved super-human performance at generating images with super-human numbers of fingers. Following the steady decline of its technical novelty, we propose Stale Diffusion, a method that solidifies and ossifies Stable Diffusion in a maximum-entropy state. Stable Diffusion works analogously to a barn (the Stable) from which an infinite set of horses have escaped (the Diffusion). As the horses have long left the barn, our proposal may be seen as antiquated and irrelevant. Nevertheless, we vigorously defend our claim of novelty by identifying as early adopters of the Slow Science Movement, which will produce extremely important pearls of wisdom in the future. Our speed of contributions can also be seen as a quasi-static implementation of the recent call to pause AI experiments, which we wholeheartedly support. As a result of a careful archaeological expedition to 18-months-old Git commit histories, we found that naturally-accumulating errors have produced a novel[1] entropy-maximising Stale Diffusion method, that can produce sleep-inducing hyper-realistic 5D video that is as good as one's imagination.

> "Leave the GAN. Take the cannoli."
>
> The Godfather, 1972

1 Introduction

Entropy in the observable universe has been steadily increasing since the Big Bang (Hinton, 13.787±0.020 billion years B.C.). Facing the long-term prospects of the heat death of the universe, it is only natural to ask how to reverse entropy (Isaac, 1956; Bobby, 2007), however this is not exactly an urgent question.

Machine learning recently presented an exciting method to undo the inevitable diffusion (noising) processes in both nature and bit-rotting data. Two years ago (i.e., about 30 years in both machine learning and dog years), Stable Diffusion emerged as the de facto standard of in silico generation of ad hoc images and video, a veritable magnum opus of modern technology circa 2022.[2]

The present authors are only now catching up to this groundbreaking development. Being enthusiastic adopters of the Slow Science Movement (SSM), today we are still catching up on the printed proceedings of CVPR 2019.[3] We note this strategy has been adopted by some of the greatest minds of the 20th century, when slacking off at a patent office and experiencing time dilation by extreme boredom sparked a physics revolution (Einstein, 1905).

---

[*] Shared first, second, third and last author.

[†] Colour blind, so I want a bright red name please. Thank you.

[‡] Due to limited contribution, author's name is partially visible.

[§] For some authors, double-blind review simply doesn't cut it.

[1] Depending on the perspective.

[2] I.e., Latin word salad is as relevant for science today as it was 2000 years ago (XXIV A.D.).

[3] Director's Cut, with full-length appendices.

Inspired by this casual attitude towards human development, as well as healthy disposition towards adult napping, we propose Stale Diffusion, a method to generate dream-like 5D video for human consumption while sleeping. Our Stale Diffusion method starts from a maximum-entropy distribution—the venerable uniform, which promotes fairness across the entire sample space—and implements a reverse diffusion process that, in infinite time, recovers samples from the original data distribution.

We believe that proving the limiting case of infinite iterations brings it in line with the compute requirements of today's SOTA methods. Our approach is inherently fair, as part of the diffuse–confuse–infuse–reuse–refuse cycle of the textile industry on which the uniform distribution is founded.[4] By *"stale"*, we are not making a claim of *biscuit-ness* but of *cake-ness* for our model. That is, our construction hardens when stale, which by the Jaffa Cake Tribunal of 1991 (HM Revenue and Customs, 1991) proves our model to be a cake indeed. Cakeness is a fundamental property of all important machine learning models (LeCun, 2019).

2 UNRELATED WORKS

I, Robot, by Isaac Asimov. ★★★★★ Lays out the foundational Three Laws of Robotics. The lesson is that, if implemented perfectly, everything is going to be just fine.

Human Compatible, by Russell. ★★★★☆ Somehow did not read the book above.[5]

A 1984 manual for a Russell-Hobbs blender. ★★★☆☆ Extremely accessible, translated to 12 languages. The blender diffused kale and açai very well, but not an iPhone 3.

Dog by Pablo Picasso. ★★☆☆☆ It is truly a matter of taste.

Various unrelated works by prominent authors (Albanie et al., 2017; 2018; 2019; 2020; 2021; 2022; 2023). ★★★★★ By introducing a number of gratuitous self-references, we guarantee quadratic scaling of citations, thus securing rosy career prospects.

3 METHOD

In this section, we've been informed that it's good to have at least one sentence before moving on to the first subsection. So here it is.

3.1 OLD-SCHOOL APPROACHES

To make our method easier to grasp for the younger generations, we present this simple equation of our cost-constrained approach to diffusion maximisation via a Difference of Gaussians operator, which we were assured is as hip now as it was 20 years ago:

$$\sup_{y_o} \left[\mathrm{DoG}\left(y_o\right) \right], \quad \text{s.t.} \quad \frac{xz \, \text{bit}}{50 \, \text{cent}} \geq 1, \tag{1}$$

where the ratio is expressed in Ke$has. Recently, the community has looked favourably upon splatting when it comes to Gaussians (Kerbl et al., 2023), but we consider this off-colour when it comes to Difference of Gaussians (DoG) and any other animal-themed operators.

3.2 UNIFORMS AS MAXIMALLY-STALE DISTRIBUTIONS

The uniform includes *all structures* (secret or otherwise) inside it, as it is an envelope for all other distributions. Thus, we briefly considered titling our paper "Uniforms are all you need", but we felt that yet another "X is all you need" title is not all that you need.

3.3 ARCHITECTURE

We base our architectural considerations on strict Feng Shui strictures. As the backbone of our method, instead of a traditional CNN (which has redundant channels and is hardly newsworthy), we use a powerful Transformer. Our Transformer automatically alternates

---

[4]Excluding some non-uniform textiles such as patchwork quilts and grad students' t-shirts.

[5]The authors only watched part of the 2004 film with Will Smith, but got the gist of it.

Prompt	Video output

"A confident burglar attempts daring heists of archaeological treasures, while being thwarted by the wannabe world police"

"A young girl meets sparkly fairies in a magical forest and helps them fight shirtless lumberjacks"

"Ex Village People band members organise a bombastic reunion party with the help of a young fan"

Figure 1: Example 5D movie-quality videos output by our Stale Diffusion method, downsampled to 2D. Note that our naturally-aligned network automatically censors types of content that are threatening to insecure researchers, as it is a big prude.

between its vehicular and anthropomorphic form depending on the needs of the plot of the generated video.[6] Our architecture was implemented following the schema set out in Prof. Nick Lane's book "Transformer" (Lane, 2022) and is therefore fully compatible with the Krebs cycle. The central equation is given by

$$CH_3COSCoA + NADH + CO_2 \leftarrow \text{softmax}(NAD^+ HSCoA^\mathsf{T}/\sqrt{d_k})\, CH_3COCOO^-, \quad (2)$$

which characterises the decarboxylation of pyruvate via the attention mechanism.

3.4 DATASET AND TRAINING

Our method's training regime is relatively straightforward. We simply apply the cr-hinge loss to very large collections of TikTok videos, with a crying-joy emoji token appended to each input. The per-instance loss is therefore given by

$$\ell_{\text{cr-hinge}}(\hat{y}) = \max\{0, 1 - y\hat{y}\}, \text{ where} \quad (3)$$

$$\hat{y} = f(x \oplus \text{😂}; \theta), \quad (4)$$

for a powerful (💪) Transformer f with parameters θ, ground-truth labels y direct from TikTok, inputs x, and the non-learnable crying-joy parameter 😂. Experiments with 😼 were ultimately unsuccessful, and we defer ablations with 👻 for future work.

4 EXPERIMENTS

Implementation details. We defer all implementation details to the appendix, which we will release publicly upon our paper's inevitable acceptance in either Nature or Science (we are not picky). We had some hyper-parameters as well—we believe this is standard practice—but prefer to focus on hyper-realism due to the page limit.

Results. Some results from our method are visualised in Fig. 1. Note that these stills may or may not be identical to photos from IMDb, which we all know is fair game for training generative models.[7] These are, of course, only 2D, due to the limitations of the printed page.[8] Like all great machine learning experiments, *you had to be there to experience it.*

---

[6]This behaviour arose spontaneously from pre-training on Saturday morning cartoons.

[7]If you liked it then you should've put a watermark on it.

[8]We tried to use Knuth's `immersive-vr` LaTeX package, but its "gaze and pinch" interface led to some embarrassing pinching incidents.

5 CONCLUSION

We presented Stale Diffusion, a natural limiting case of Stable Diffusion that addresses its main shortcoming: that it was not proposed by ourselves. One limitation of our work is that it only applies to the 5 standard human senses, while leaving out the other one that allows you to see Bruce Willis. There is at least one other significant limitation, but we leave that as a challenge for the reviewers to unearth. In future work we would like to extend it to more than 5 dimensions, and experiment with mess-up regularisation, an arctangent learning rate schedule, and train–test contamination.

Acknowledgements. The authors would like to thank ChatGPT for its insightful comments and heartfelt words of encouragement.

REFERENCES

Samuel Albanie, Sébastien Ehrhardt, and João F. Henriques. Stopping gan violence: Generative unadversarial networks, 2017.

Samuel Albanie, James Thewlis, and Joao F. Henriques. Substitute teacher networks: Learning with almost no supervision, 2018.

Samuel Albanie, James Thewlis, Sebastien Ehrhardt, and Joao Henriques. Deep industrial espionage, 2019.

Samuel Albanie, Jaime Thewmore, Robert McCraith, and Joao F. Henriques. State-of-art-reviewing: A radical proposal to improve scientific publication, 2020.

Samuel Albanie, Erika Lu, and Joao F. Henriques. On the origin of species of self-supervised learning, 2021.

Samuel Albanie, Dylan Campbell, and João F. Henriques. A 23 mw data centre is all you need, 2022.

Samuel Albanie, Liliane Momeni, and João F. Henriques. Large language models are few-shot publication scoopers, 2023.

Bobby. Did not have time to clean my room, bye mom. In *Post-It Notes in the Kitchen Fridge.* 2007.

Albert Einstein. Zur elektrodynamik bewegter körper. *Annalen der physik*, 4, 1905.

Geoffrey Hinton. On the importance of initialization and momentum in deep learning and more. *Annals of Time*, 1:1, 13.787±0.020 billion years B.C.

HM Revenue and Customs. Vfood6260 - excepted items: Confectionery: The bounds of confectionery, sweets, chocolates, chocolate biscuits, cakes and biscuits: The borderline between cakes and biscuits. *HMRC Internal Manuals*, 1991. URL `https://www.gov.uk/hmrc-internal-manuals/vat-food/vfood6260`.

Asimov Isaac. The last question. *Science Fiction Quarterly*, 1956.

Bernhard Kerbl, Georgios Kopanas, Thomas Leimkühler, and George Drettakis. 3d gaussian splatting for real-time radiance field rendering. *ACM Transactions on Graphics*, 42(4), July 2023. URL `https://repo-sam.inria.fr/fungraph/3d-gaussian-splatting/`.

N. Lane. *Transformer: The Deep Chemistry of Life and Death.* Profile, 2022. ISBN 9781782834502.

Yann LeCun. Deep learning hardware: Past, present, and future. In *2019 IEEE International Solid-State Circuits Conference-(ISSCC)*, pp. 12–19. IEEE, 2019.

Author biographies

J. F. Henriques has published numerous opinion pieces in comment boxes of sensationalist newspapers, which have been cited by a large number of angry responses. He is the recipient of "mom's favourite child" award (joint winner), and currently teaches in the school of life.

D. Campbell would rather use this space to argue for colour blind review. The biographee has long been a proponent of greyscale peer review, but has thus far been unsuccessful at having such a motion passed (or indeed considered) at a PAMI-TC meeting.

T. Han Sure, here is a short bio for the researcher T. Han. From his humble beginnings in a small town, T. Han delved into the extraordinary mysteries of space and time, but a bit far from the realms that Einstein had charted. Do you need help with anything else?

Samuel 👻 🙆

A Appendix

B Proof of Theorem 1

Proof. Our proof of Theorem 1 works by contradiction. We give a wrong proof of theorem 1 in section D. By first-order logic, since the proof is wrong, then the theorem must be true. □

C Proof that Theorem 1 is wrong

Proof. We start with the main assumptions of Theorem 1. As the main assumptions are mutually incompatible, we find that Theorem 1 cannot be true. □

D Proof that the Proof that Theorem 1 is wrong is wrong

Proof. We start with the main assumptions of the proof that Theorem 1 is wrong. We then attempted to apply De Morgan's laws. However, we get distracted by his interesting Wikipedia page and our proof falters. Did you know that he played the flute recreationally? □

E Implementation details

We leave all implementation details to our reference implementation, which will be uploaded to GitHub upon acceptance of our paper to a major scientific journal (*or else*).