

Towards Generative Geometric AI

Generative AI for 360° panoramic images



Jason McEwen · Following

5 min read · Apr 19

Listen

Share

More



AI generated 360° image. [Original image generated by author using [CopernicAI](#).]

Generative AI has made remarkable progress in recent years, enabling machines to generate images, text, and even music. However, a number of data modalities are still missing.

Some of the most notable generative models include [GPT-4](#), a language model that can generate human-like text, and [DALL-E 2](#), an image generation model that can create high-quality images from textual descriptions. [Stable Diffusion](#) has made significant strides in generating high-quality images. As research continues to progress, the possibilities for generative AI seem almost limitless.

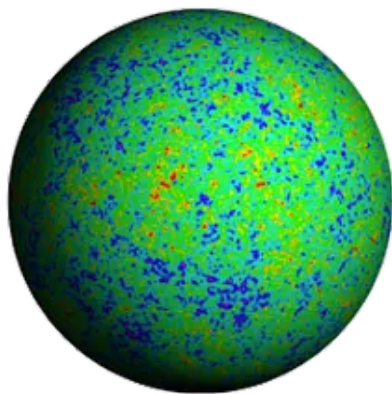


Photograph of an astronaut riding a horse generated by Stable Diffusion. [Image sourced from [Wikimedia Commons](#).]

Limitations

Most generative AI models, however, are tailored to data defined on Euclidean domains, such as 2D images or 1D audio signals. Nevertheless, in many fields, data are often defined on non-Euclidean geometries, giving rise to the field of geometric deep learning (see [*A Brief Introduction to Geometric Deep Learning*](#)).

Approaches to geometric deep learning are broad but one of the most common cases is the group setting (see [*Geometric Deep Learning on Groups*](#)), such as observations defined on the sphere. Spherical data arise in many fields, from observations over the Earth to panoramic 360° photos and videos in virtual reality, to astronomical observations of the relic light from the Big Bang, which is observed on the celestial sphere.



Cosmic microwave background



360° virtual reality

Examples of spherical data. [Original figure created by author.]

To bring the benefits of generative AI to data like these and others with complex geometries we need to merge geometric and generative AI.

Generative AI architectures

Modern generative AI models typically take one of the following approaches.

1. *GANs* (Generative Adversarial Networks) use a generator and a discriminator network to generate new data samples by pitting the two against each other in a zero-sum game.
2. *VAEs* (Variational Autoencoders) are generative models that learn a compressed representation of the input data and use it to generate new samples.
3. *Normalizing flows* are a class of generative models that use a series of invertible transformations to model the probability distribution of the data.
4. *Diffusion models* use a denoising diffusion process to generate high-quality images.

The current state-of-the-art in terms of generated image quality is typically achieved by GANs or diffusion models. We, therefore, concentrate on these approaches and explore how they may be extended to geometric settings.

Towards generative geometric AI for 360° images

For concreteness, let us focus on the common case of spherical data and in particular, 360° panoramic images.



Little planet view of a 360° panoramic image. [Image sourced from [Wikimedia Commons](#).]

While both normalizing flows and diffusion models have been extended to the spherical setting in recent pioneering research papers [1,2], these approaches focus on a density field over spherical coordinates, whereas here, we are interested in a pixelated image on the sphere.

To extend both GANs and diffusion models to spherical data, the underlying architectures of these approaches must be extended to natively model the underlying spherical geometry. To achieve this, we require the underlying layers that form the building blocks of this architecture to be extended to the sphere. In the past, such spherical layers have been plagued by computational limitations. Recently, however, these computational limitations have been overcome in the hybrid discrete-continuous (DISCO) framework [3], which provides spherical layers that are both highly effective and computationally efficient (see our recent article on [*Hybrid Discrete-Continuous Geometric Deep Learning*](#)).

GANs are built on a discriminator (i.e., classifier) and a generator that supports dense predictions, often themselves built on CNN layers for image data (see [*A Brief Introduction To GANs*](#) for further details). All of these components have been extended to the sphere already [3,4]; hence, we have all of the building blocks needed to extend GANs to spherical data.

Diffusion models require a learned score or, equivalently, a denoiser, often based on U-Net style architectures (see [*An Introduction to Stable Diffusion*](#) for further details). Again, U-Nets for high-resolution images have already been extended to the sphere [3], and thus we have precisely the building blocks that we need to handle.

In the meantime

At [CopernicAI](#), we are working on precisely these types of architectures to bring generative AI to geometric data, such as 360° images.

Until these new models are ready, however, we have retrofitted Stable Diffusion to support the generation of 360° images.

Stable Diffusion does not correctly model the underlying spherical nature of 360° data, so what can be achieved by retrofitting is somewhat limited and the quality of generated images suffers a little. Nevertheless, we can already generate fairly good 360° images.



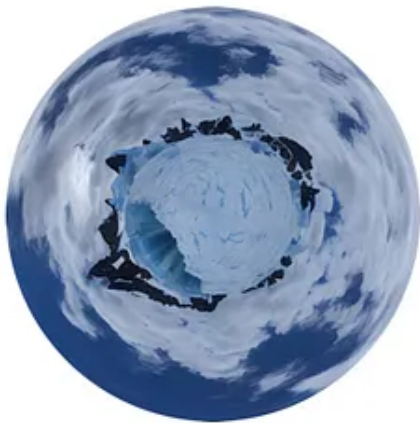
“Stonehenge”



“Home in Atlantas”



“Egyptian Desert”



“Antartica”



“Japanese Spa”



“Park in City”



AI-generated 360° images shown as little planets. [Original images generated by the author using CopernicAI.]

A [prompt available here](#) has just been released by CopernicAI so that anyone can now create AI-generated 360° images for their use.

And with future developments where the underlying geometry of the sphere is the correct model, as described above, the quality of generated 360° images is only set to increase.

References

[1] Rezende, Papamakarios, Racanière, Albergo, Kanwar, Shanahan, Cranmer, *Normalizing Flows on Tori and Spheres*, ICML (2020), [arXiv:2002.02428](#)

[2] De Bortoli, Mathieu, Hutchinson, Thornton, Teh, Doucet, *Riemannian Score-Based Generative Modelling*, NeurIPS (2022), [arXiv:2202.02763](#)

[3] Ocampo, Price, McEwen, *Scalable and equivariant spherical CNNs by discrete-continuous (DISCO) convolutions*, ICLR (2023), [arXiv:2209.13603](#)

[4] McEwen, Wallis, Mavor-Parker, *Scattering Networks on the Sphere for Scalable and Rotationally Equivariant Spherical CNNs*, ICLR (2022), [arXiv:2102.02828](#)