

Обзор на статью DreamFusion: Text-to-3D using 2D Diffusion

Кратко суть статьи

- Хотим генерировать 3D модели не используя 3D данные в обучении.
- Для генерации используем NeRF.
- Обучаем его с помощью 2D диффузионной модели и видов 3D модели с разных ракурсов.

Авторы



Ben Pool
Google Research



Ajay Jain
UC Berkeley



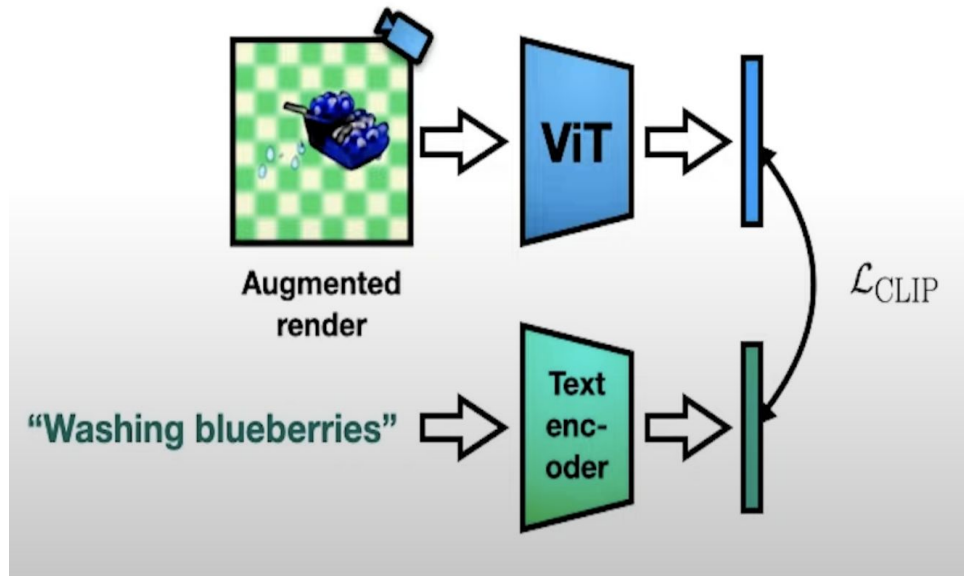
Jon Barron
Google Research



Ben Mildenhall
Google Research

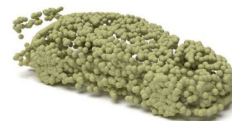
Авторы

До этого авторы совместно работали над Dream Fields. Суть статьи та же, только вместо диффузионной модели используем CLIP.



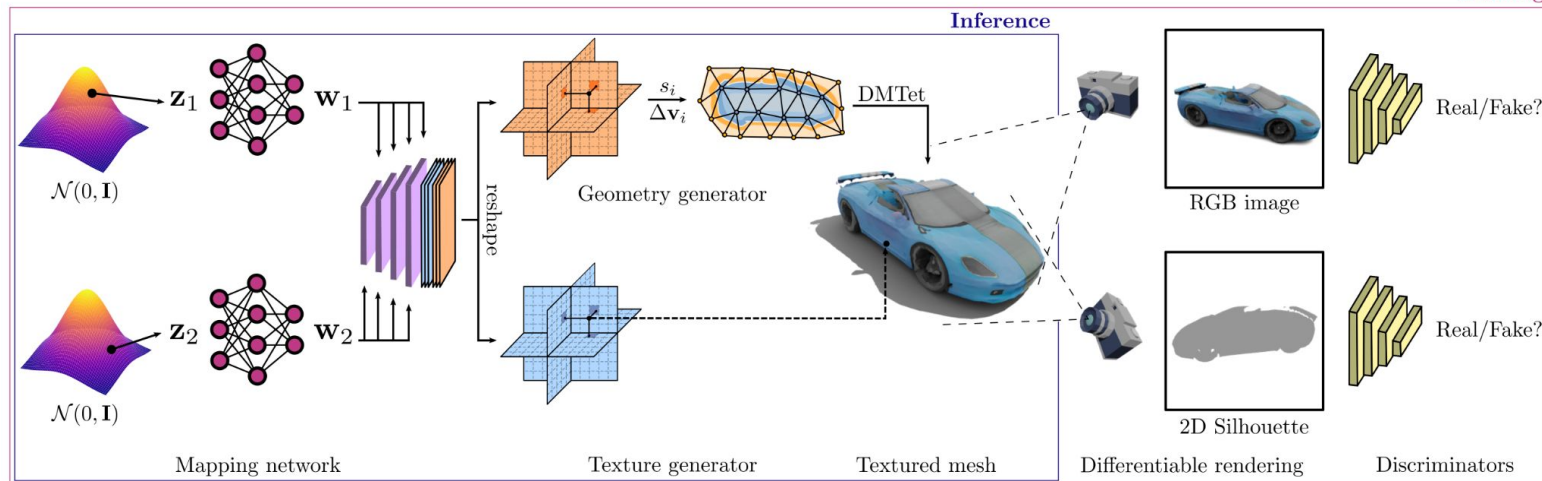
Предшественники

- NeRF
- Модели использующие воксели
- Модели использующие облака точек
- GANы
- DreamFields



Связанные статьи. Get3D

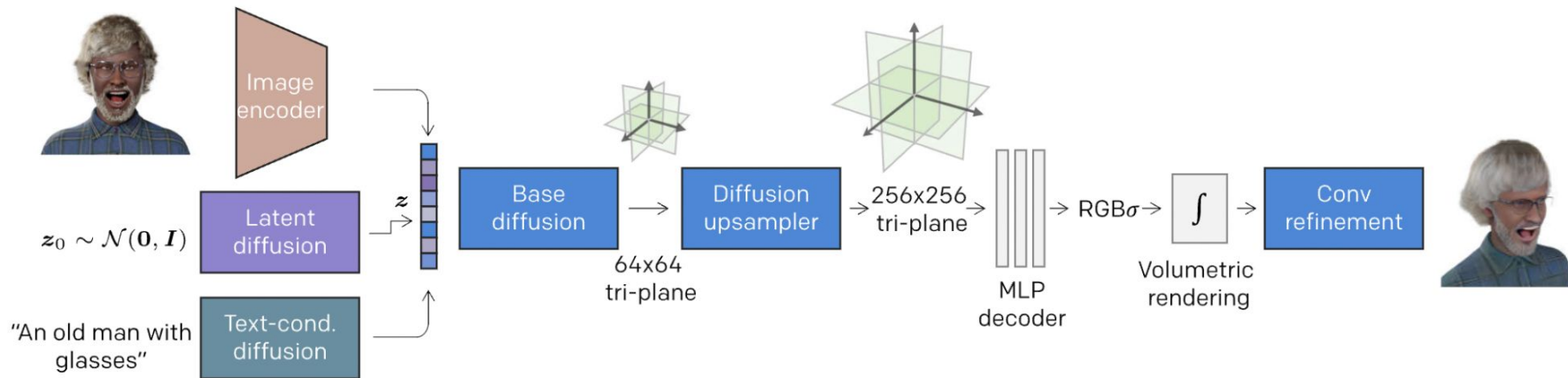
Training



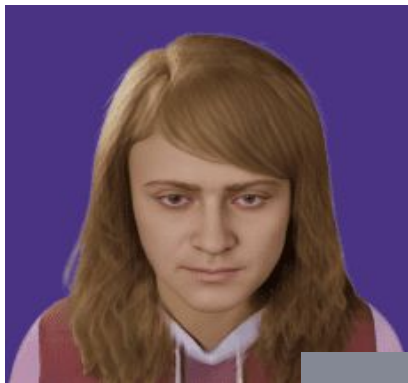
Связанные статьи. Get3D



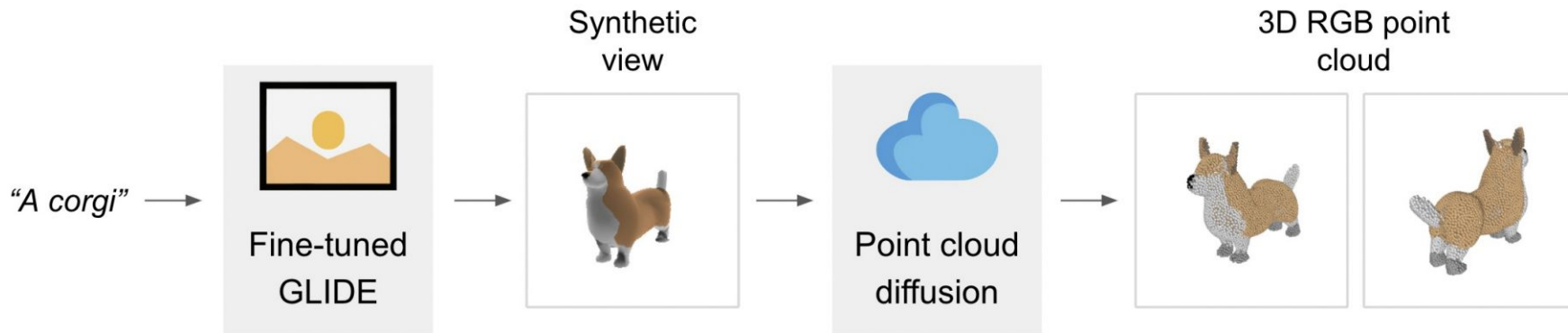
Связанные статьи. 3D Avatar Diffusion



Связанные статьи. 3D Avatar Diffusion



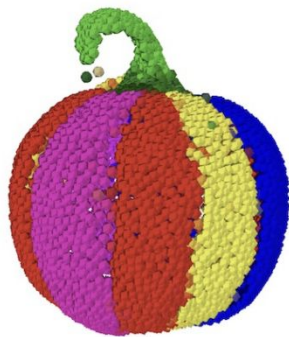
Связанные статьи. Point-E



Связанные статьи. Point-E



“a corgi wearing a
red santa hat”



“a multicolored rainbow
pumpkin”

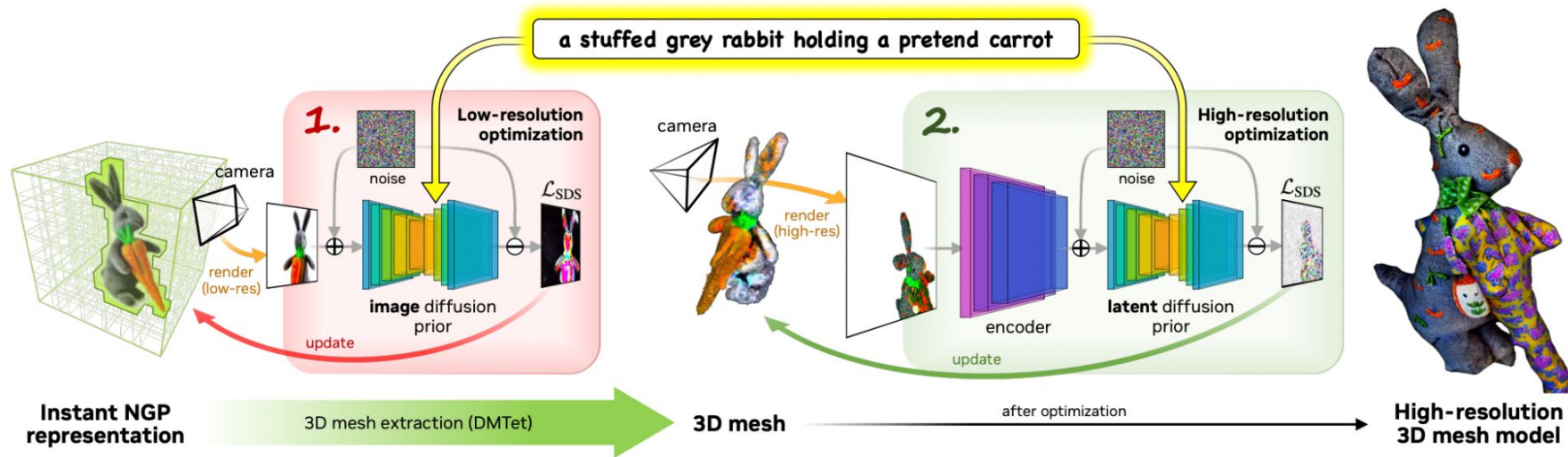


“an elaborate fountain”

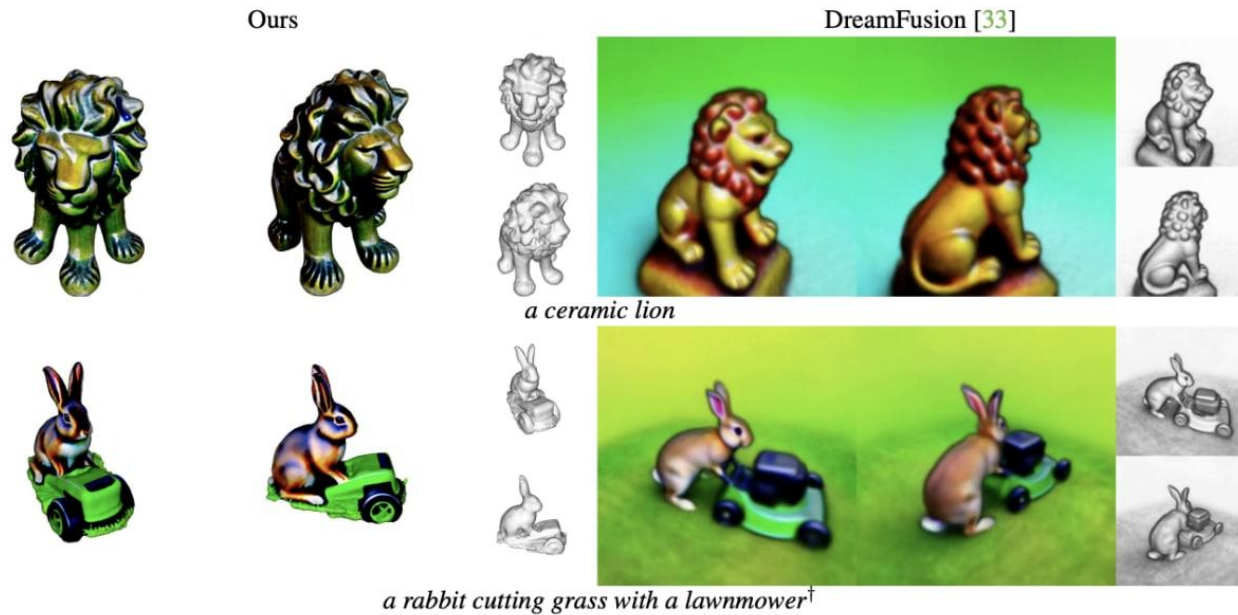


“a traffic cone”

Связанные статьи. Magic3D



Связанные статьи. Magic3D



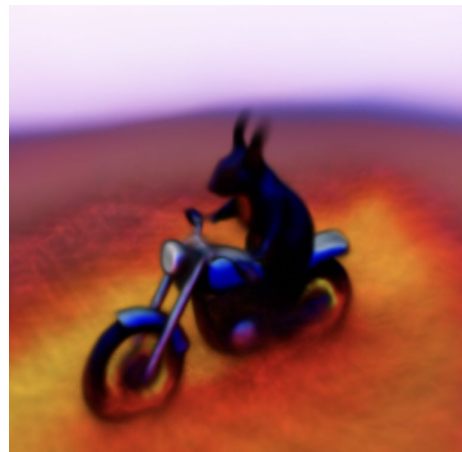
Плюсы статьи

- Нужна только 2D диффузионная модель (!!!)
- Можно получить разнообразные объекты.
- Можно контролировать Guidance.
- Придуман новый лосс, который можно использовать в других задачах.



Недостатки статьи

- Картинки получаются размытые, бывают проблемы с освещением.
- Иногда у объекта появляется 2 лица.
- Модель очень тяжелая и 3D модели генерируются больше часа.
- Диффузионная модель плохо учитывает направление камеры.
- Проведено очень мало экспериментов. В качестве бейзлайна только DreamFields и одна другая модель.



Итоги

Идея генерировать 3D модели используя модели для 2D кажется мне очень перспективной. Остается только работать над скоростью генерации и качеством получаемых моделей.



Бонус

Compare DreamFusion, Magic3d and Point-E.

✓ Searching for: **compare DreamFusion Magic3d Point-E**

✓ Searching for: **Magic3d**

✓ Generating answers for you...

I found some information about DreamFusion, Magic3D and Point-E for you. 😊

DreamFusion, Magic3D and Point-E are three different methods of generating **3D models from text prompts** ^{1 2 3 4 5 6 7 8 9}. They are all based on **neural networks** and **Neural Radiance Fields (NeRF)**, which allow a neural network to store 3D models ^{1 2 3 4 5 6 7 8 9}.

DreamFusion was developed by a team from **Google Research and UC Berkeley** and presented in a paper in **September 2022** ^{1 4}. It uses a pretrained 2D text-to-image diffusion model called **Imagen** to generate 2D images from text, and then uses NeRF to reconstruct 3D models from the images ^{1 4 6}.

Magic3D was developed by a team from **Nvidia** and presented in a paper in **November 2022** ^{5 6 7 8 9}. It uses a **coarse-to-fine strategy** leveraging both low- and high-resolution diffusion priors to learn the 3D representation of the target content ^{5 6 7 8 9}. It also supports **image conditioning techniques** and **prompt-based editing approach** to provide users with more control over 3D synthesis ^{5 6 7 8 9}.

Point-E was developed by a team from **OpenAI** and presented in a paper in **December 2022** ^{2 3 4}. It uses a pretrained text-to-image model called **DALL-E** to generate 2D images from text, and then uses NeRF to reconstruct 3D models from the images ^{2 3 4}. However, unlike DreamFusion and Magic3D, Point-E does not generate 3D models in the classical sense, but **point clouds** representing 3D shapes ^{1 2 3 4}.

Some of the main differences among the three methods are:

- **Quality:** Magic3D can create high-quality 3D mesh models with **8× higher-resolution supervision** than DreamFusion, while also being **2× faster** ^{5 6 7 8 9}. User studies show **61.7%** raters to prefer Magic3D over DreamFusion ^{5 6 7 8 9}. Point-E has a **comparatively lower quality** than both DreamFusion and Magic3D, as it generates point clouds instead of mesh models ^{1 2 3 4}.
- **Speed:** Point-E is meant to be **faster** than both DreamFusion and Magic3D, as it generates point clouds instead of mesh models ^{2 3 4}. It is **approximately 600 times quicker** to generate than DreamFusion ². Magic3D is **2× faster** than DreamFusion, as it uses a coarse-to-fine strategy to generate 3D models ^{5 6 7 8 9}.
- **Control:** Magic3D supports **image conditioning techniques** and **prompt-based editing approach** to provide users with more control over 3D synthesis ^{5 6 7 8 9}. For example, users can use an image as a reference or modify the text prompt to change the appearance of the 3D model ^{5 6 7 8 9}. DreamFusion and Point-E do not support these features, as they only use text prompts as inputs ¹.

Learn more: [1. the-decoder.com](#) [2. mythicalai.substack.com](#) [3. the-decoder.com](#) [+6 more](#)

Статья

- <https://arxiv.org/pdf/2209.14988.pdf>

