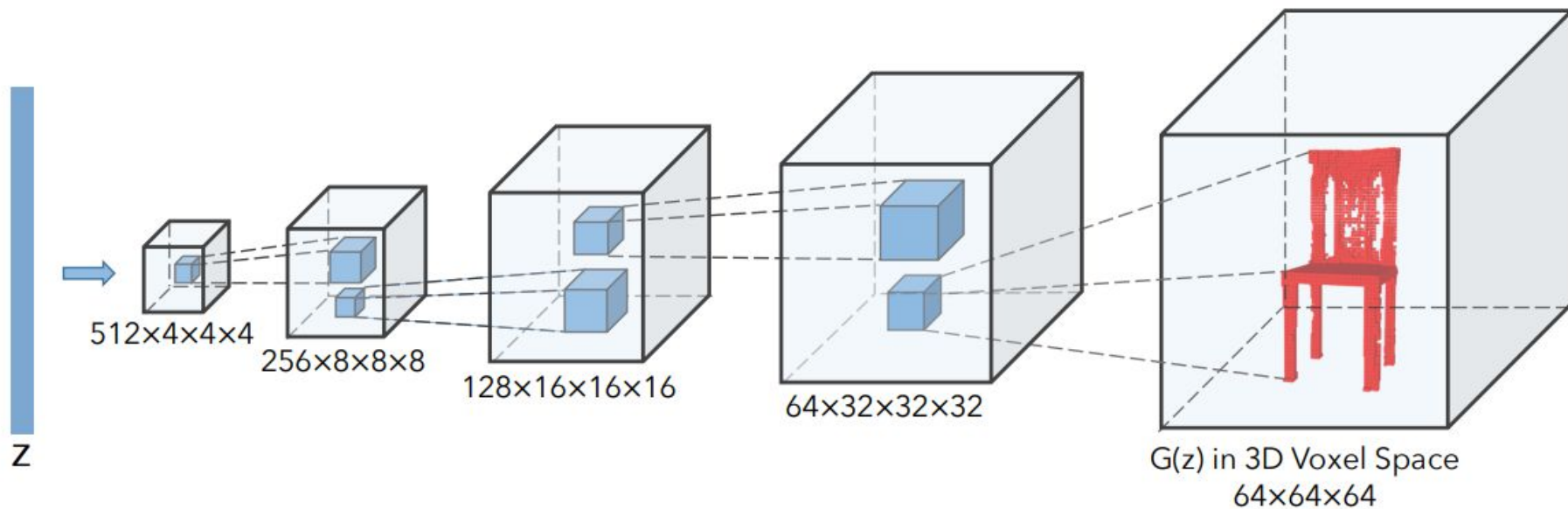


3D-Chair-GAN

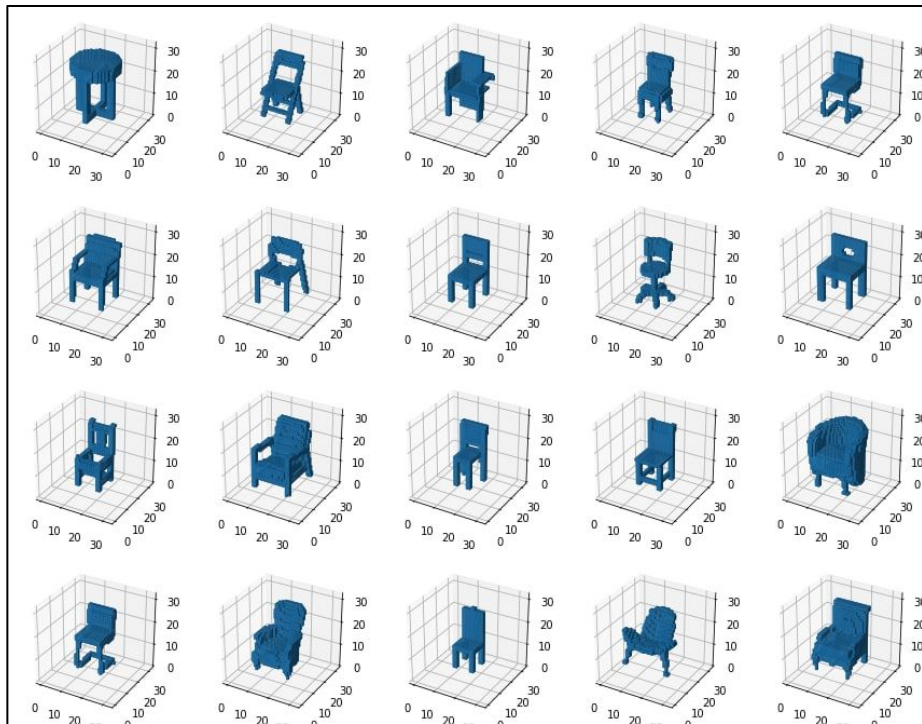
Ярослав Гераськин, Тихон Воробьев

Задача

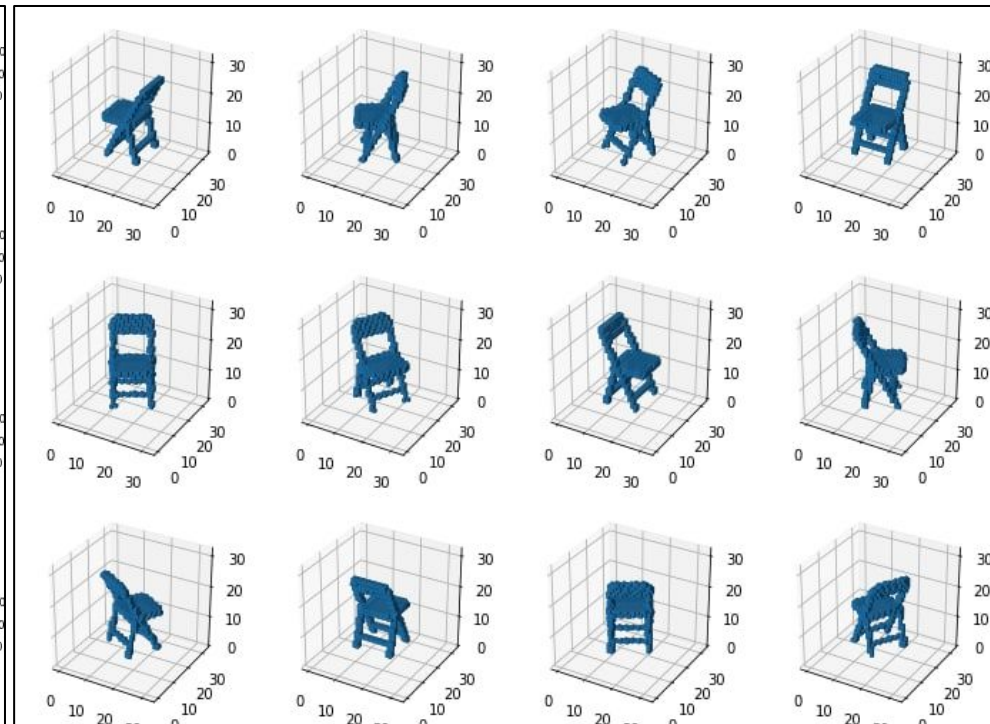
Генерация трехмерных стульев в воксельном представлении на основе результатов статьи [Learning a Probabilistic Latent Space of Object Shapes via 3D Generative-Adversarial Modeling](#).



Данные

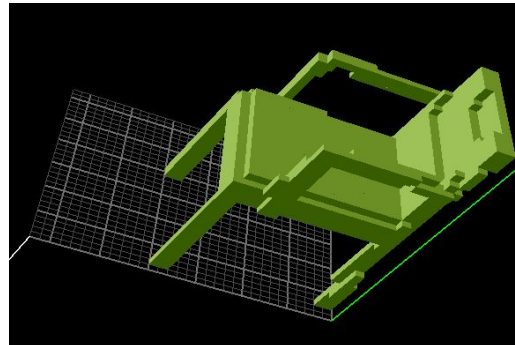
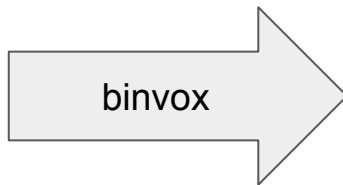
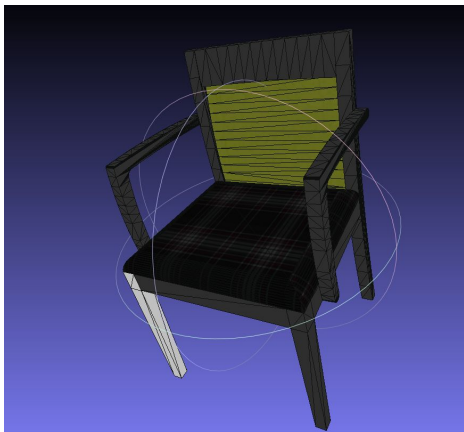


Разные стулья (889 штук)



Под разными углами (12 поворотов, всего 10668)

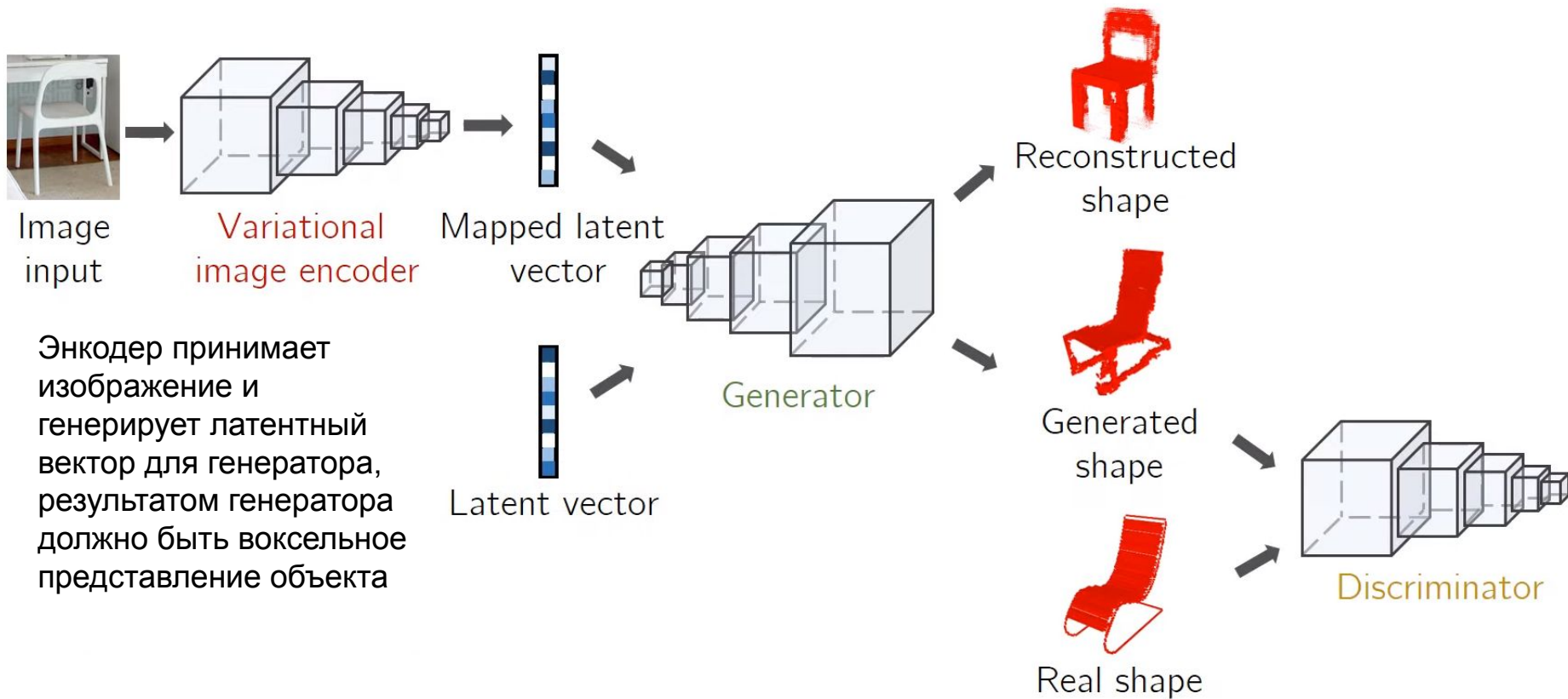
Данные (Shapenet data)



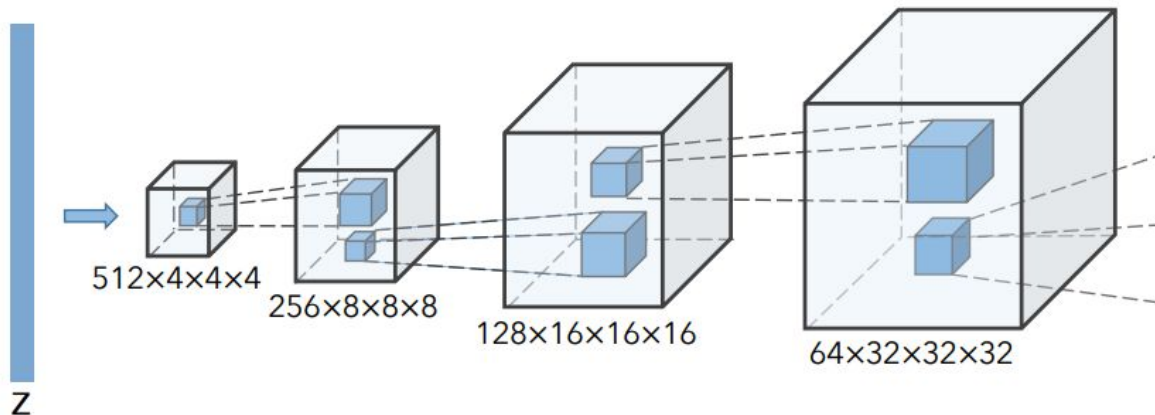
SUN
backgrounds

- 6752 стула
- По 2 рендера на каждую модель
- Воксельное представление 32x32x32 в формате binvox

VAE 3D GAN

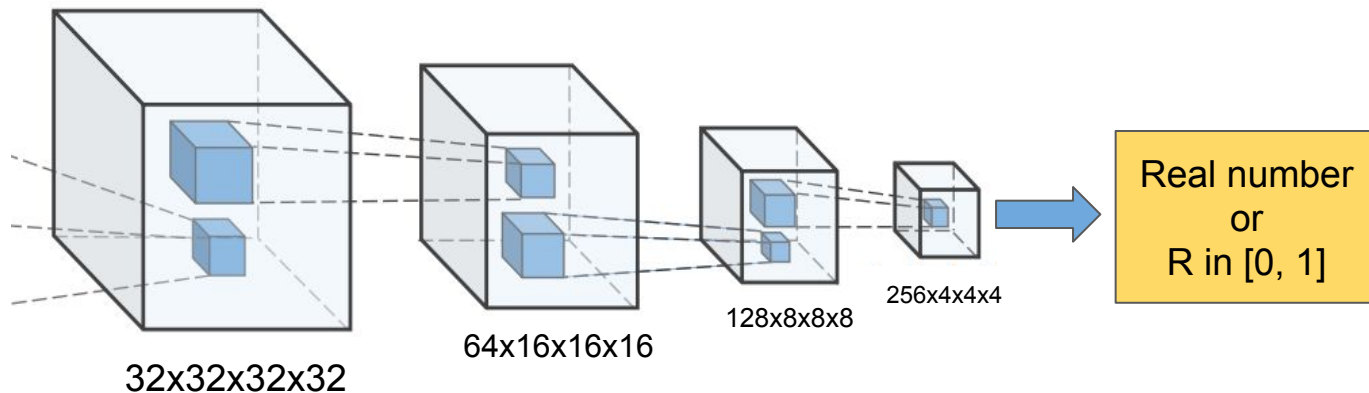


Модели (Генератор)



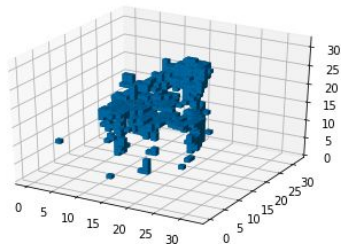
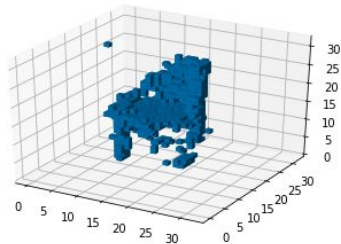
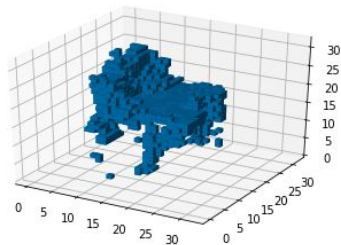
Генератор такой же, как в статье, но выход $32 \times 32 \times 32$, $\dim(z)=200$
Структура моделей такая же, как в статье, но с модификациями

Модели (Дискриминатор)

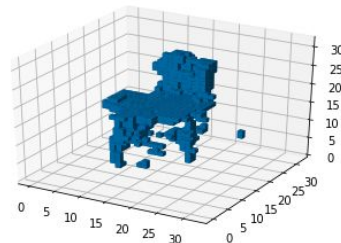
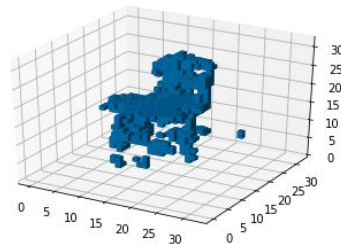
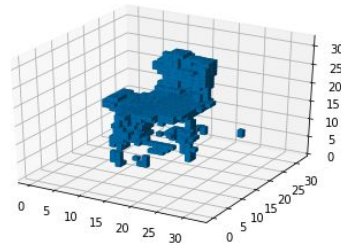


Дискриминатор такой же, как в статье
Структура моделей такая же, как в статье, но с модификациями

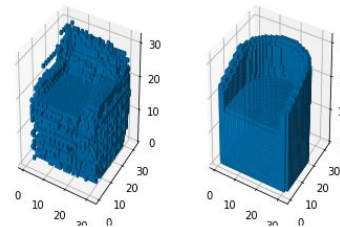
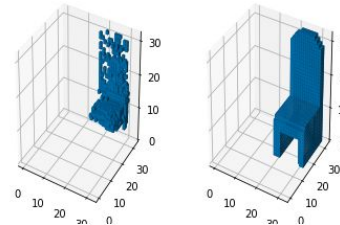
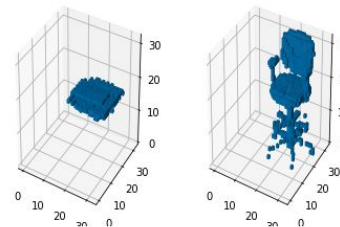
Промежуточные результаты



3D GAN



VAE 3D GAN



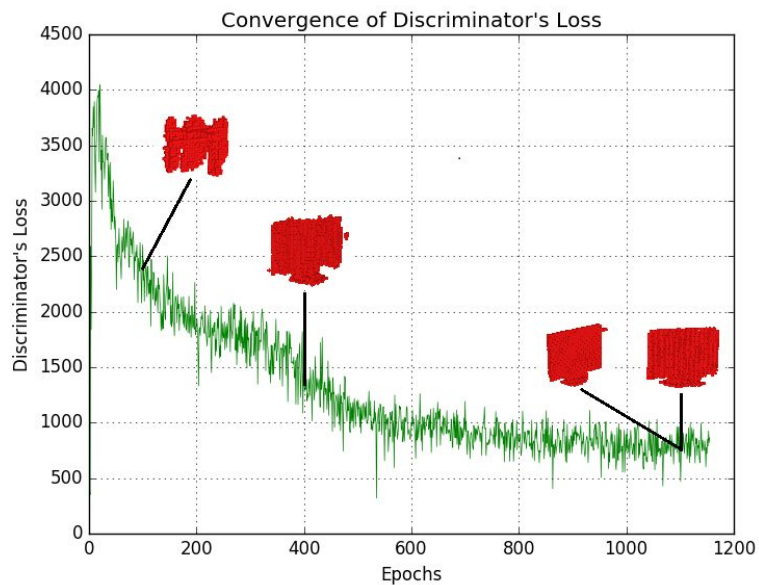
Модификации

- Gen BN -> Gen IN
- Disc BN -> Disc SN ([Spectral normalization](#))
- GAN -> Hinge WGAN ([Geometric GAN](#))
- [TTUR](#) (use different lr for gen and disc)

$$L_D = -\mathbb{E}_{(x,y) \sim p_{data}}[\min(0, -1 + D(x, y))] - \mathbb{E}_{z \sim p_z, y \sim p_{data}}[\min(0, -1 - D(G(z), y))]$$

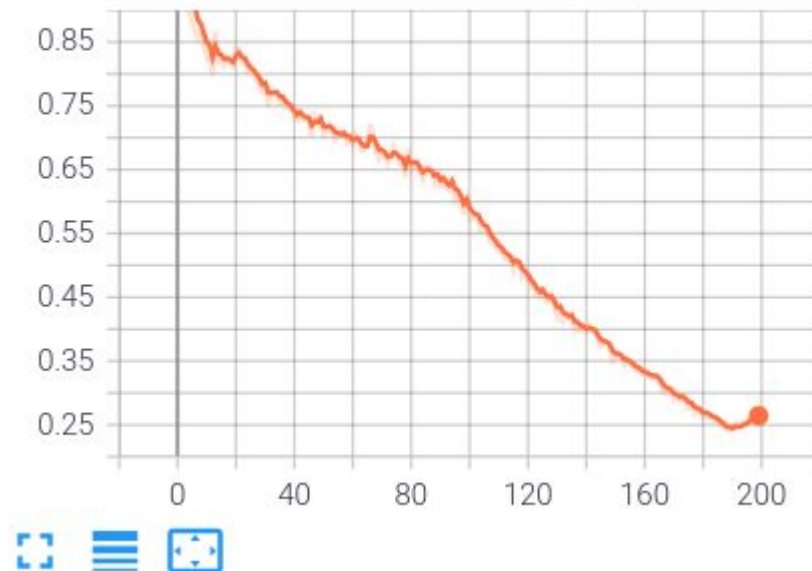
$$L_G = -\mathbb{E}_{z \sim p_z, y \sim p_{data}} D(G(z), y)$$

Обучение



3D-IWGAN

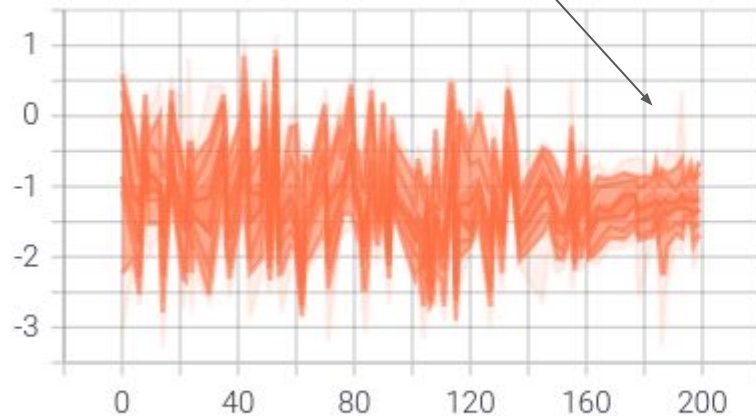
Loss/disc
tag: Loss/disc



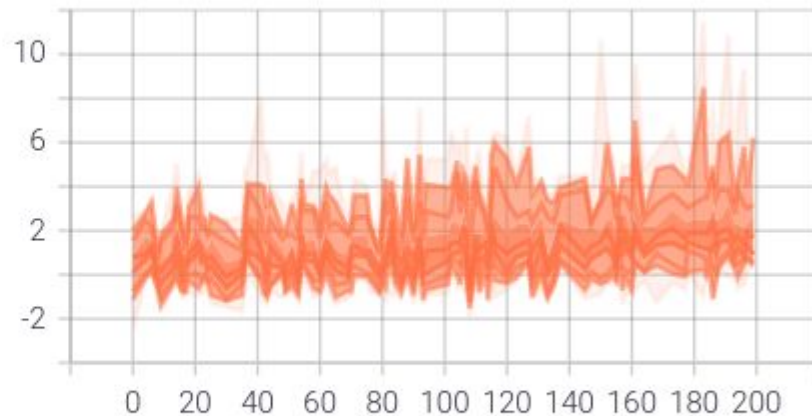
Ours

Обучение

Hist/fake
tag: Hist/fake

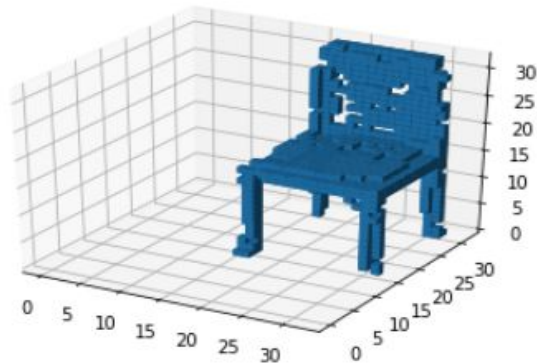
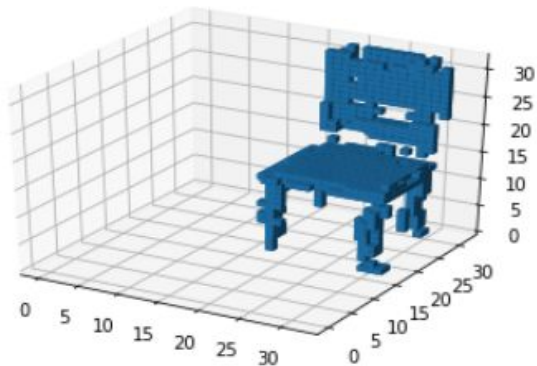
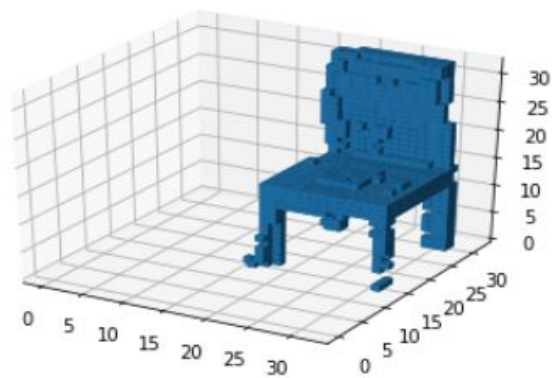
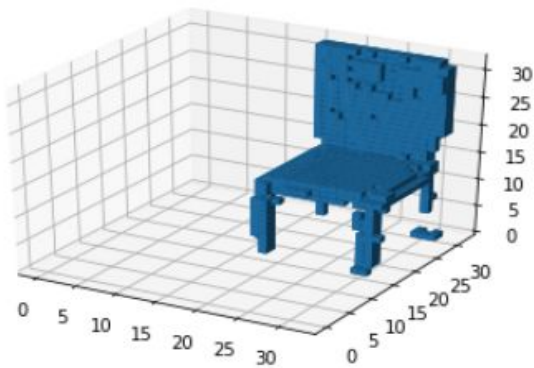


Hist/real
tag: Hist/real



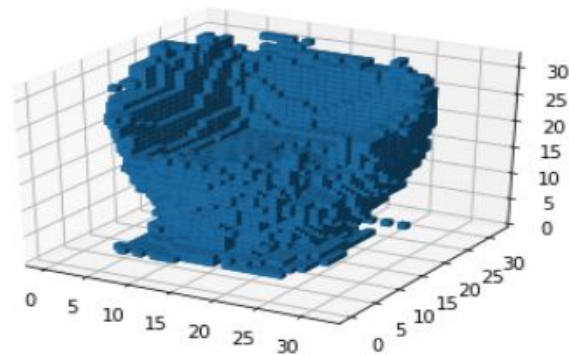
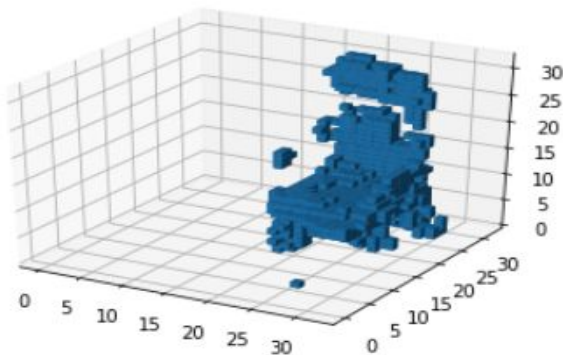
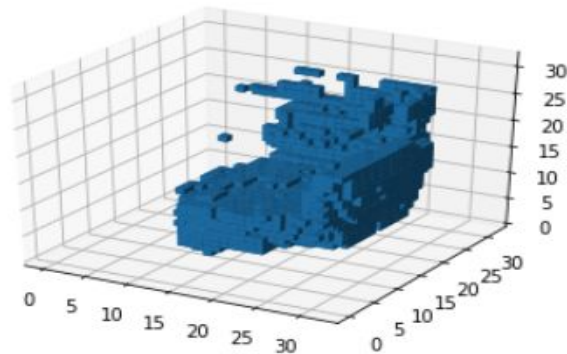
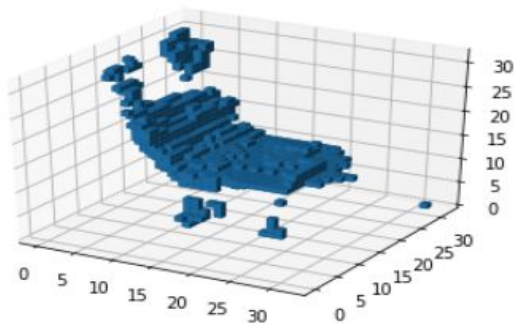
Распределение значений на выходе дискриминатора для сгенерированных объектах показывает, что сгенерированные объекты больше похожи друг на друга, чем реальные стулья

Результаты (Shapenet data)



Один вид стульев (видимо, самый простой) получается хорошо

Результаты (Shapenet data)



Другие типы стульев тоже есть, но для модели они оказались сложнее

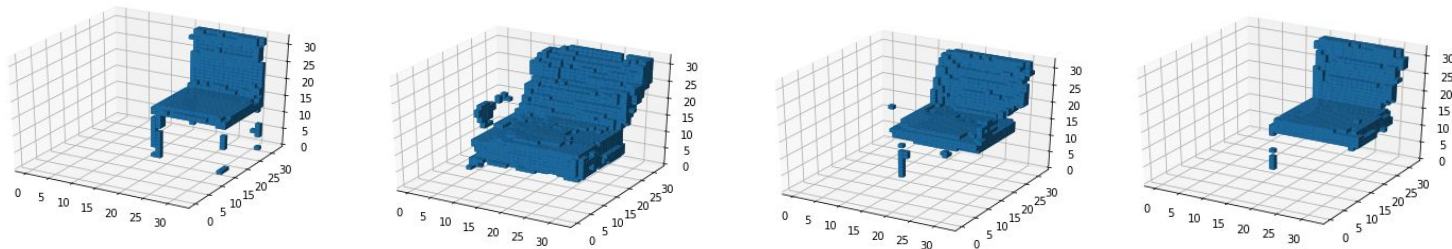
Обучение энкодера

- Энкодер обучался при фиксированном генераторе
- Только reconstruction loss (L1)

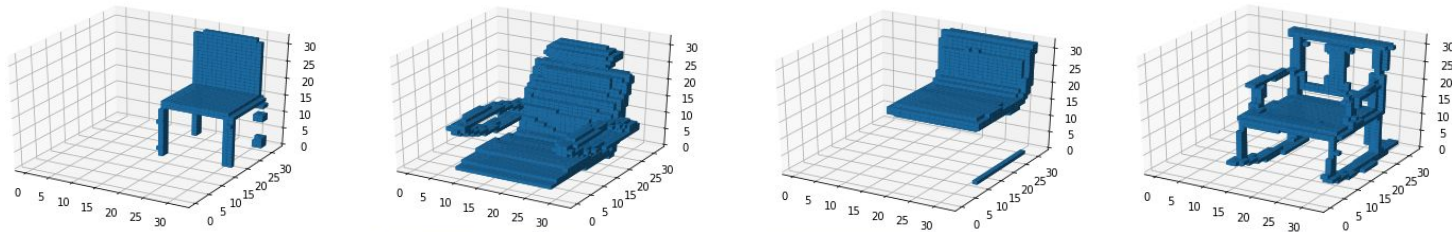
Encoder Layer	Input
conv2d 11x11 s=4 ch=32	3x256x256
conv2d 5x5 s=2 ch=64	32x64x64
conv2d 5x5 s=2 ch=128	64x32x32
conv2d 5x5 s=2 ch=256	128x16x16
conv2d 8x8 s=1 ch=2*200	256x8x8
Flatten	400x1x1
Sample	400
Output (Identity)	sample (200), mean (200), sigma (200)

Результаты (Энкодер)

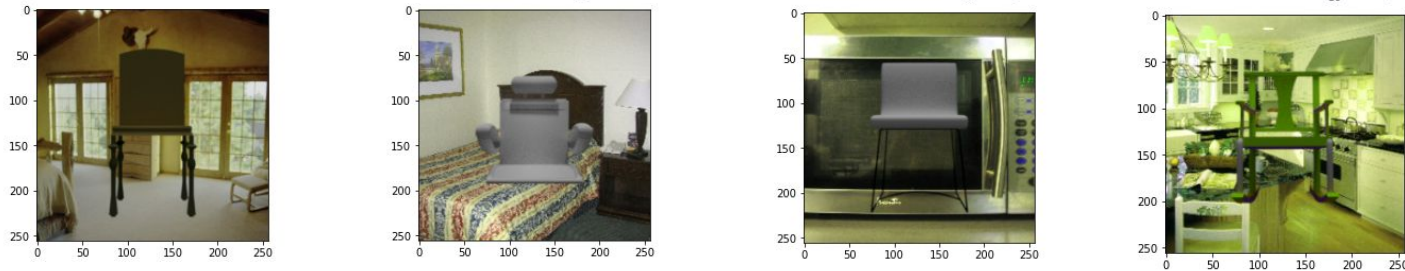
Generated model



Real model

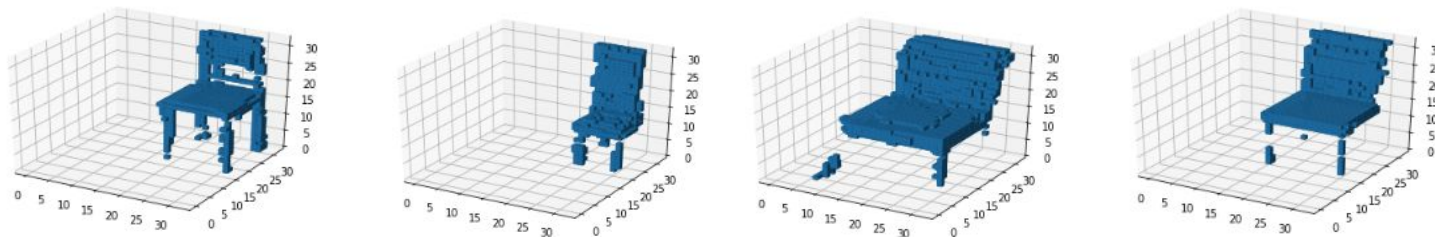


Encoder input

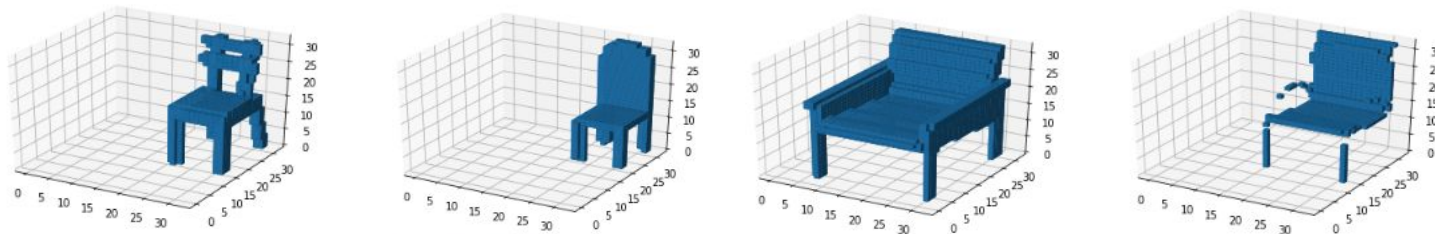


Результаты (Энкодер)

Generated model



Real model



Encoder input

