

深層学習を用いた3次元モデル 超解像の学習安定化の検討

工学研究科 情報システム工学分野

指導教員 椋木雅之

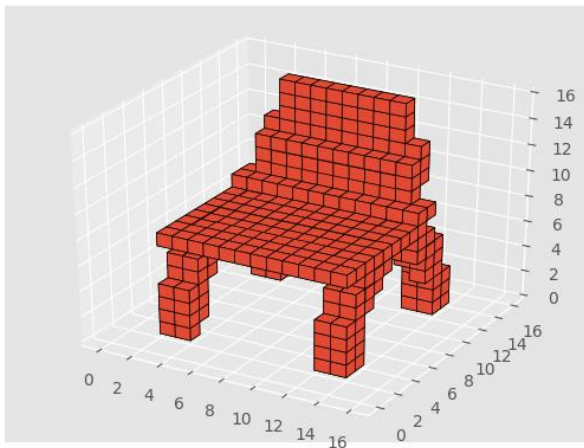
T1703047 森芳雄

2020/02/04

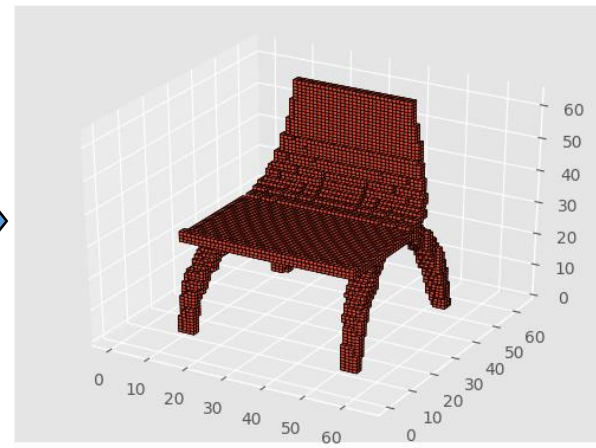
3次元モデルの超解像

3D-SRGAN

(3D-Super Resolution Generative Adversarial Networks)[1]



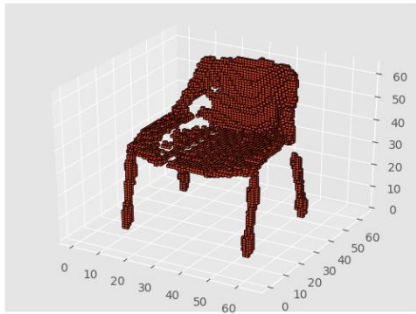
低解像度3次元モデル



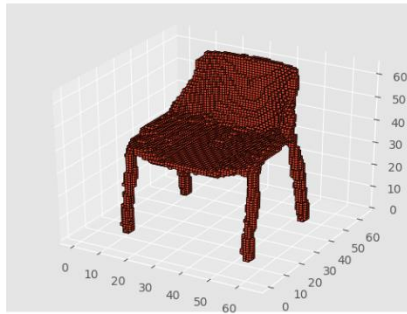
高解像度3次元モデル

[1] 岡和寿, “SRGANの3次元モデル超解像への拡張”, 平成30年度 宮崎大学 工学研究科 情報システム工学分野修士論文, 2019.

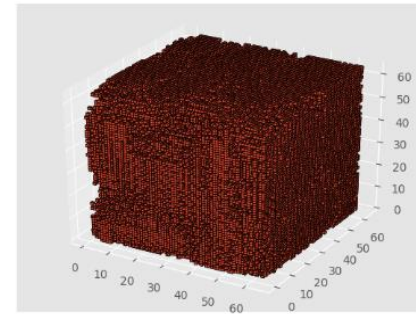
3D-SRGANの問題点



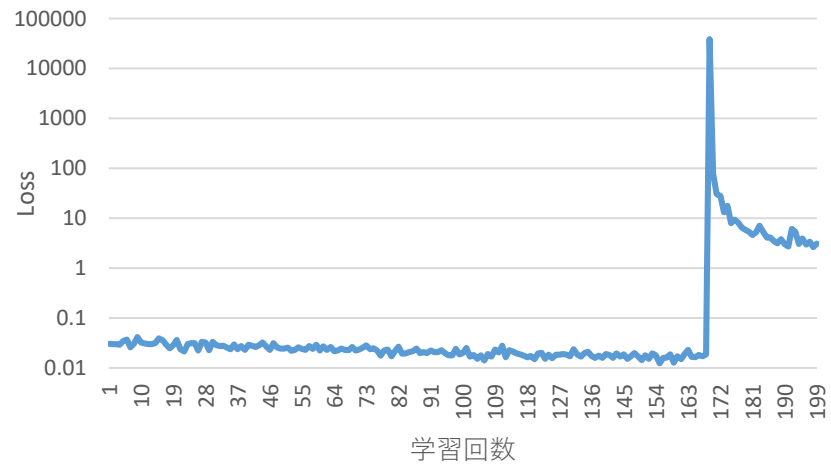
100回



150回

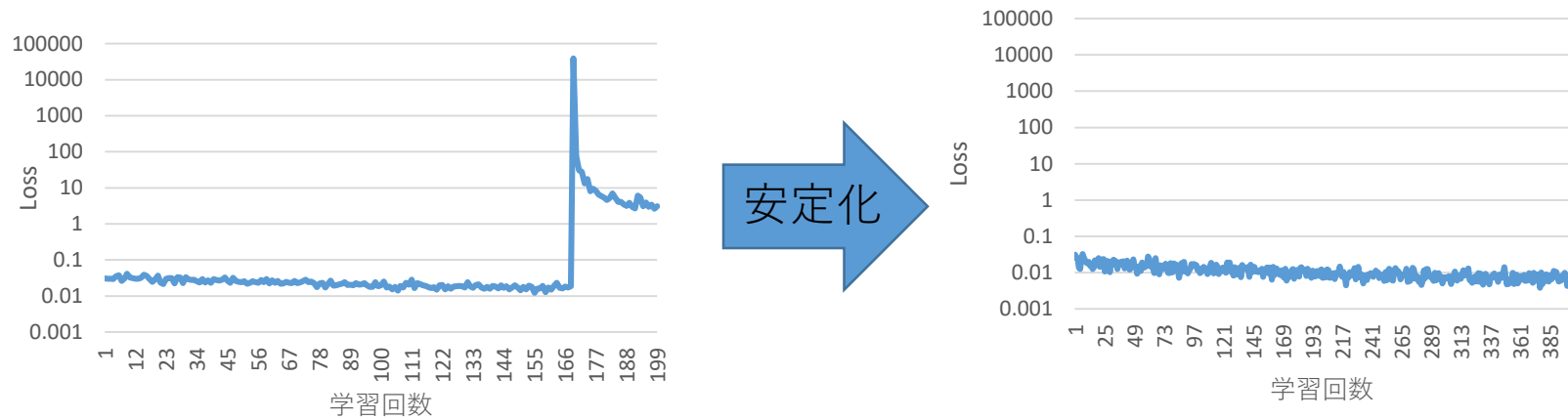


200回



研究の目的

3D-SRGANの学習の安定化について検討を行う



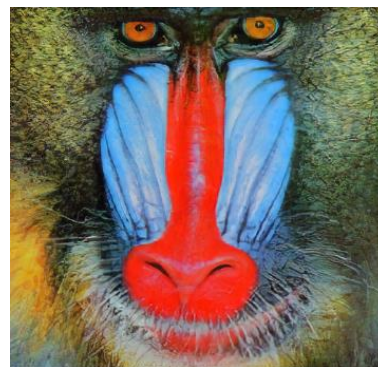
3D-SRGAN

画像の超解像で有効性が知られているSRGAN[2]を
3次元モデルが扱えるように拡張

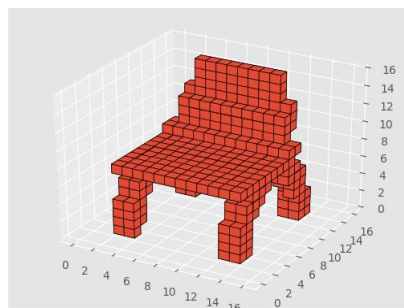
SRGAN



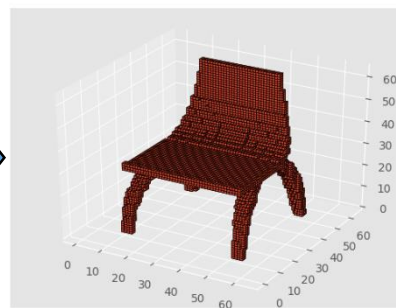
超解像



3D-SRGAN

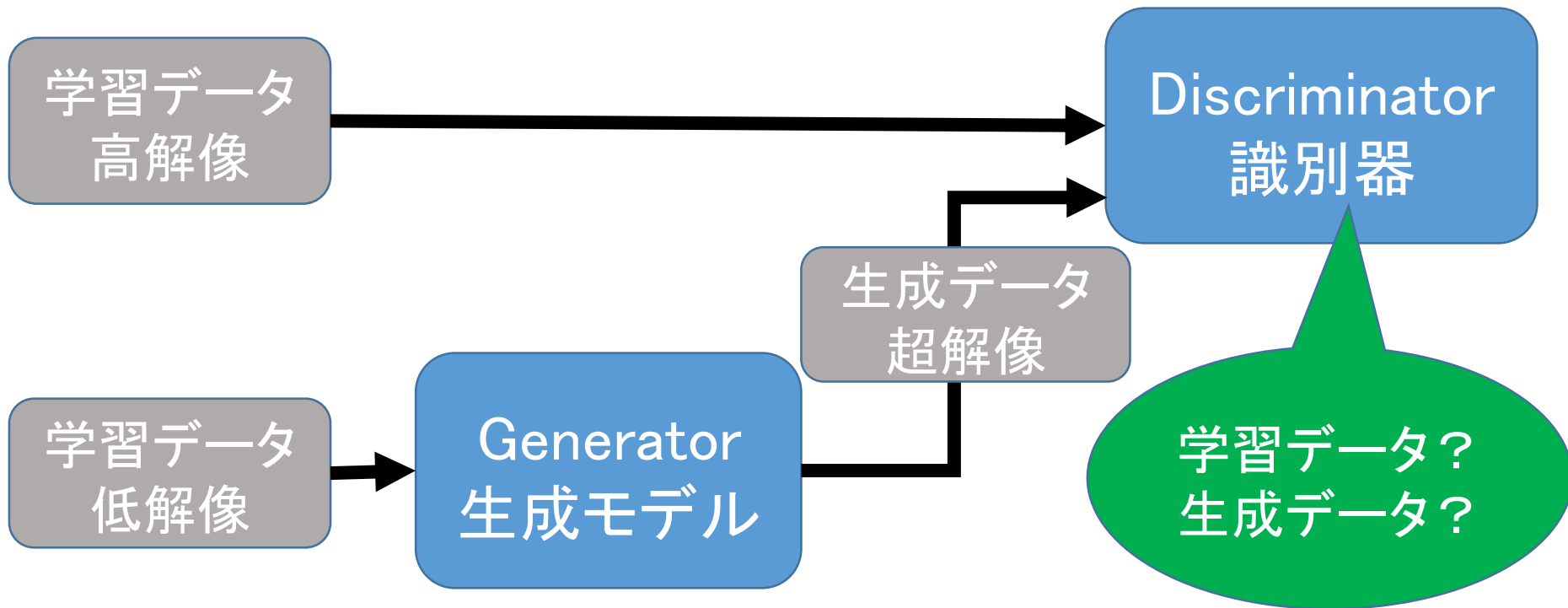


超解像



3D-SRGANの学習

Generatorが学習データに近い3次元モデルを生成できることが目的



3D-SRGANの学習

- 目的関数を最小化

$$E(w) = \mathbb{E}_{\tilde{x} \sim \mathbb{P}_g} [\log(1 - D(\tilde{x}))]$$

w : ネットワークのパラメータ

- 収束計算は勾配 $\nabla E(w)$ に基づく

$$w' \leftarrow w - \alpha F(\nabla E(w))$$

α : 学習率

学習の不安定性

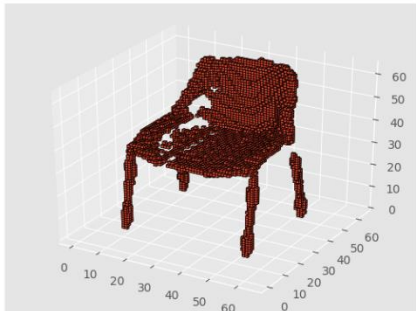
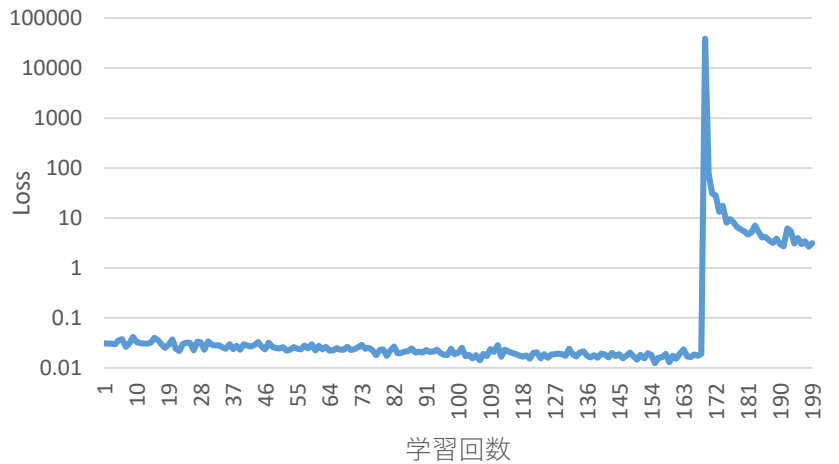
勾配消失： $\nabla E(w) \sim 0$ $w' \sim w$

勾配爆発： $\nabla E(w) \sim \text{大}$ $w' \sim \text{大}$

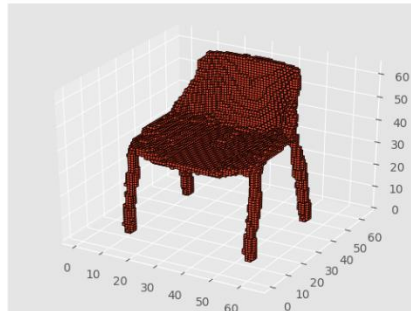
正規化などさまざまな処理を行い安定化を図ってる
→かえって問題を大きくしているかもしれない

3D-SRGANの学習の不安定さ

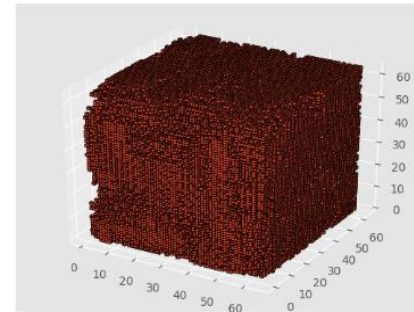
LossはGeneratorの
目的関数の収束度合



100回

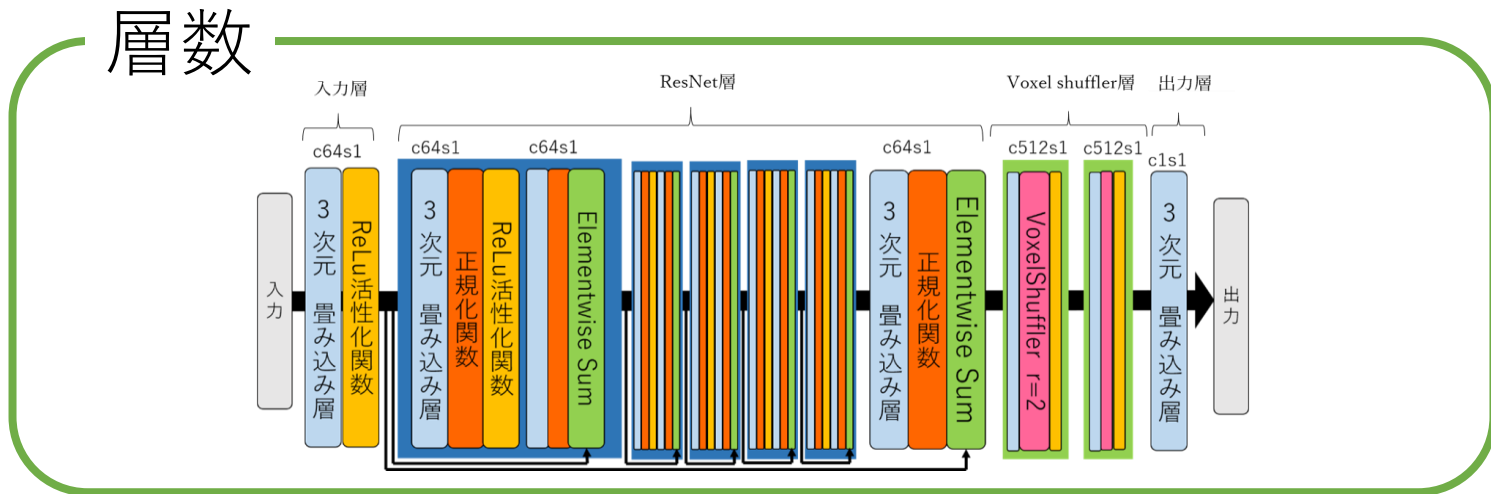


150回



200回

不安定性の原因



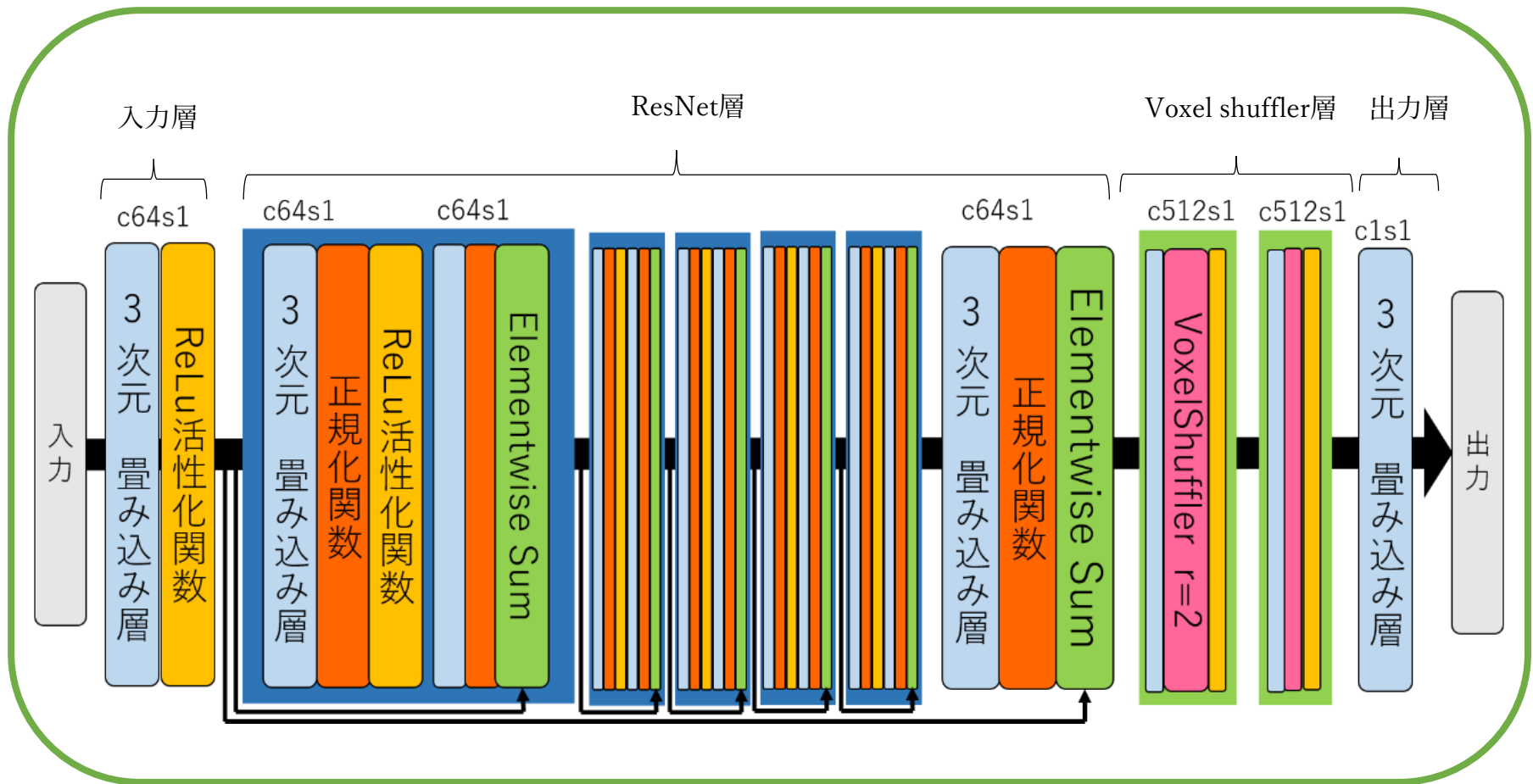
目的関数

$$\mathbb{E}_{x \sim \mathbb{P}_r} [\log D(x)] + \mathbb{E}_{\tilde{x} \sim \mathbb{P}_g} [\log(1 - D(\tilde{x}))]$$

ハイパーパラメータ

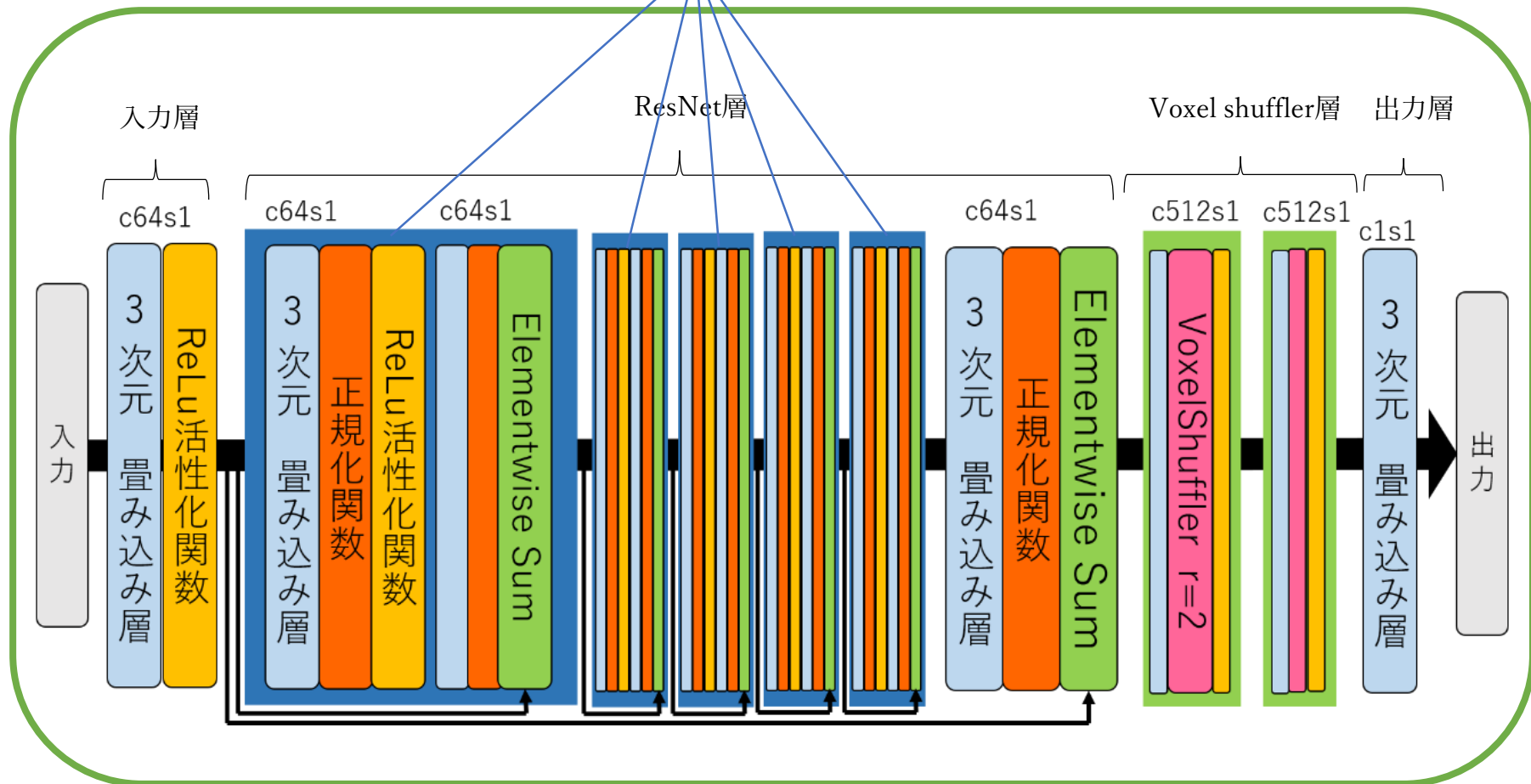
Generator vs Discriminator

層数



層数

Residual Block



目的関数

GANの目的関数

$$\mathbb{E}_{x \sim \mathbb{P}_r} [\log D(x)] + \mathbb{E}_{\tilde{x} \sim \mathbb{P}_g} [\log(1 - D(\tilde{x}))]$$



WGAN-GP[3]の目的関数

$$\mathbb{E}_{\tilde{x} \sim \mathbb{P}_g} [D(\tilde{x})] - \mathbb{E}_{x \sim \mathbb{P}_r} [D(x)] + \underbrace{\lambda \mathbb{E}_{\hat{x} \sim \mathbb{P}_{\hat{x}}} [(\|\nabla_{\hat{x}} D(\hat{x})\|_2 - 1)^2]}_{\text{Gradient Penalty}}$$

ハイパーパラメータ

ハイパーパラメータ：学習時に用いるパラメータ

Generatorの学習が過度に進むと学習が不安定になる傾向がある

学習の進み方の調整

WGAN-GP

GeneratorとDiscriminatorの更新頻度を変更する
学習率は小さい値を使う

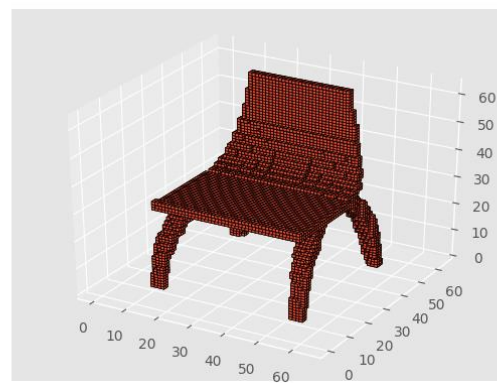
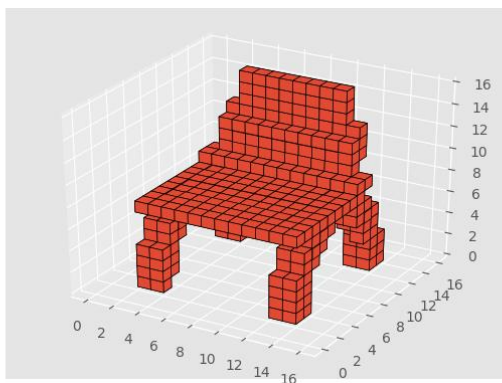
実験

実験内容

1. 層数の違いによる学習安定化の調査
2. 学習手法の変更による学習安定化の調査
3. 別クラスの超解像による精度評価

実験前準備

- ModelNet10
10クラスの3次元モデル { 学習用データ
テスト用データ
- 低解像度3次元モデル : $16 \times 16 \times 16$
- 高解像度3次元モデル : $64 \times 64 \times 64$

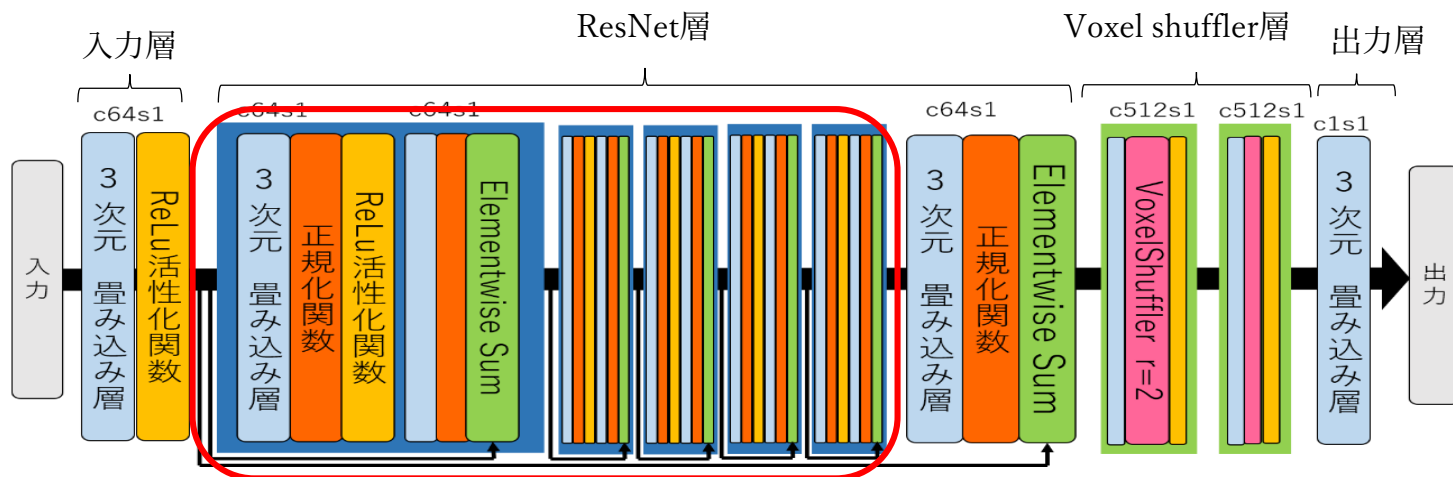


実験1：実験方法

3D-SRGANの層数を変化させることで学習の安定化につながるか調査した

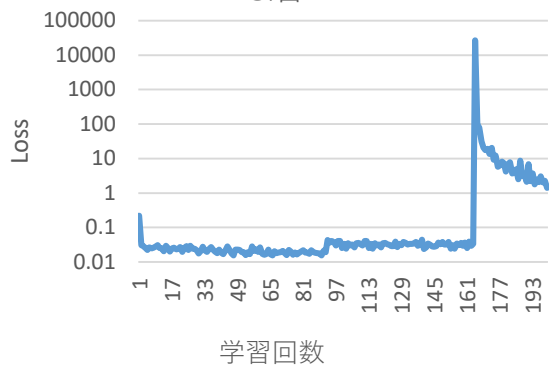
3D-SRGANのResNet層のResidual Blockの数を
変更させ学習を行った

3層、4層、5層、6層、7層

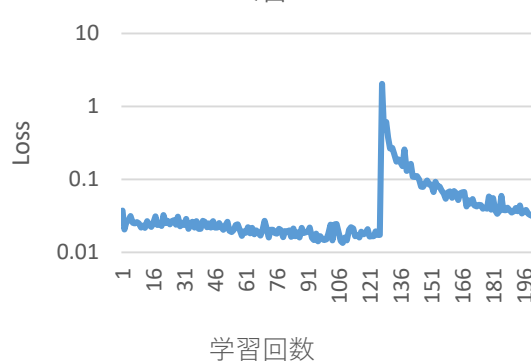


実験1：実験結果(Loss)

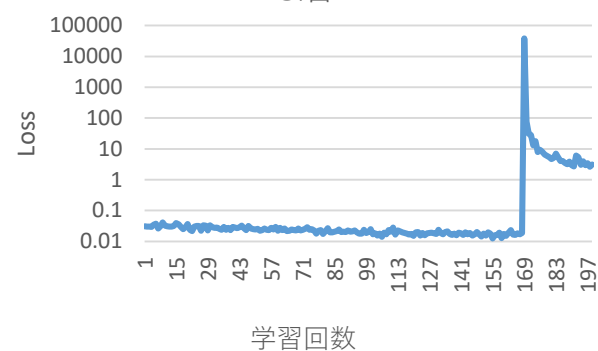
3層



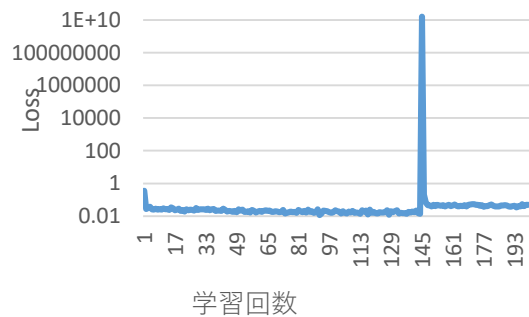
4層



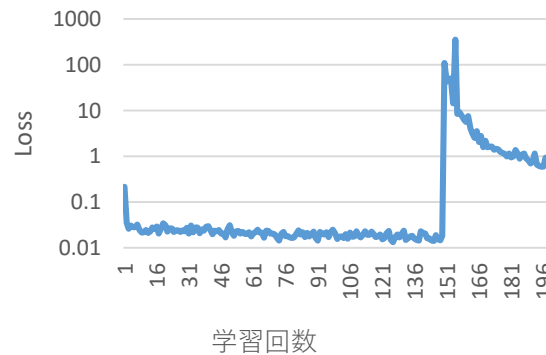
5層



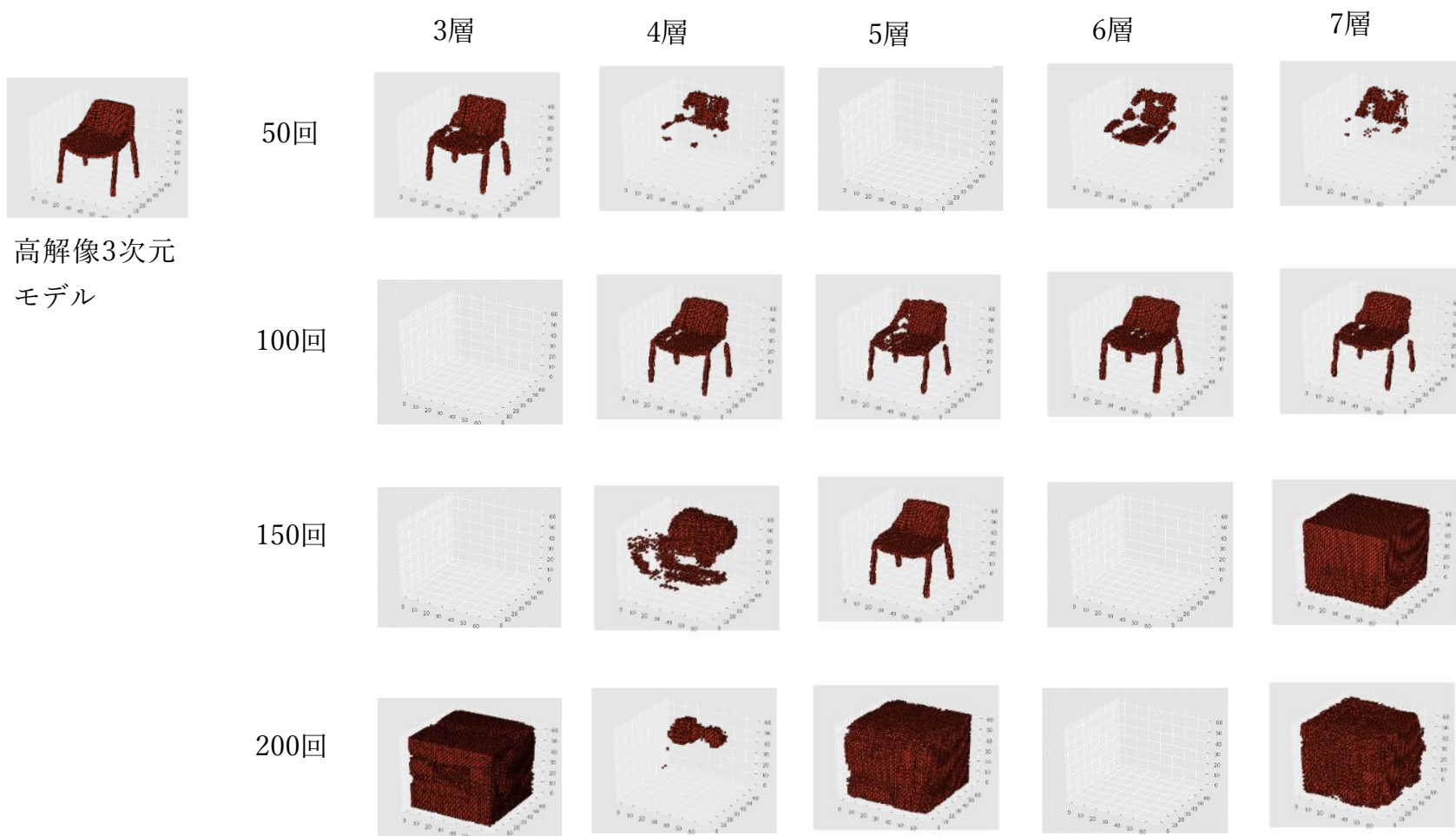
6層



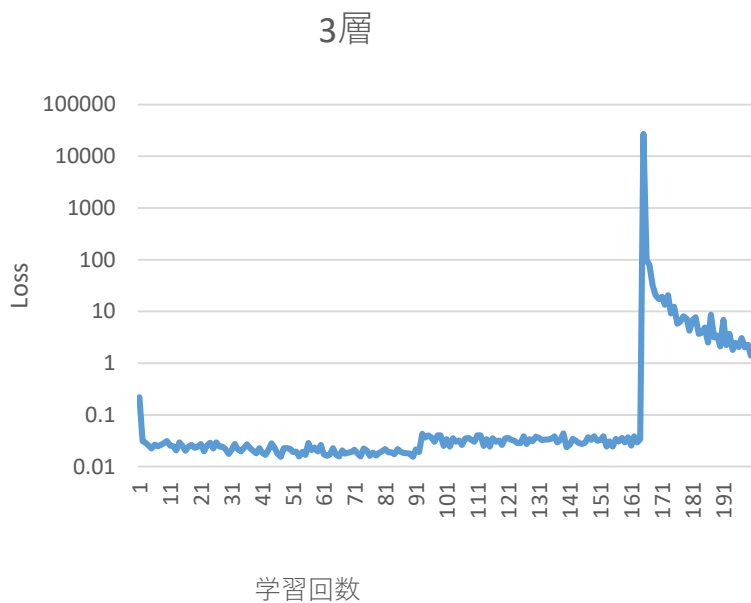
7層



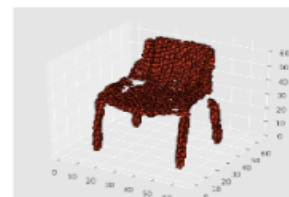
実験1：実験結果(生成結果)



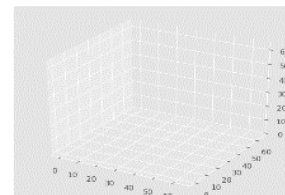
実験1：実験結果(3層)



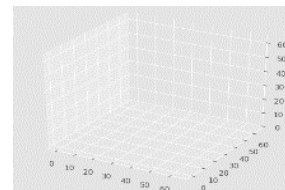
50回



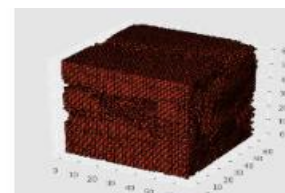
100回



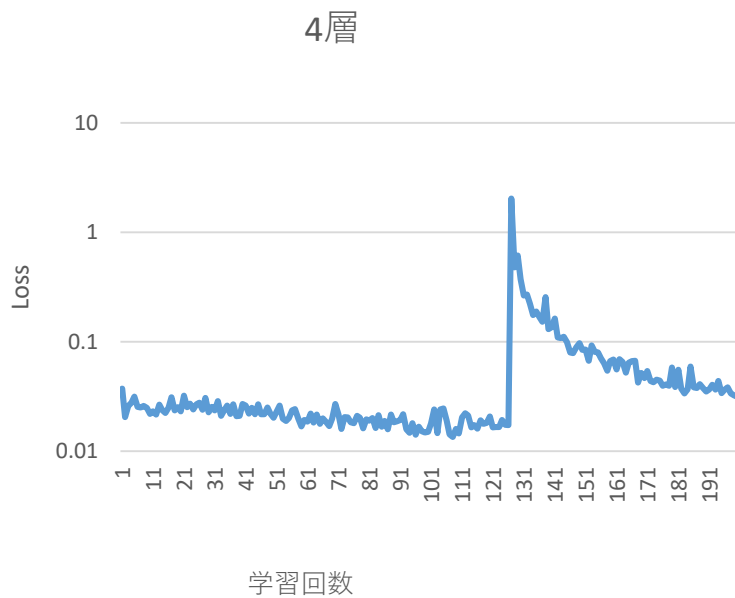
150回



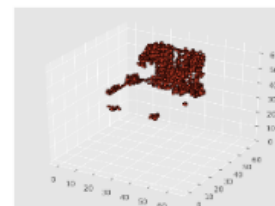
200回



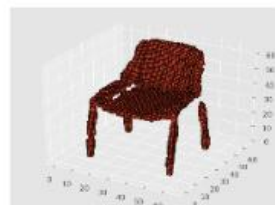
実験1：実験結果(4層)



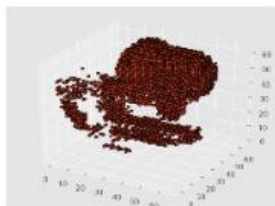
50回



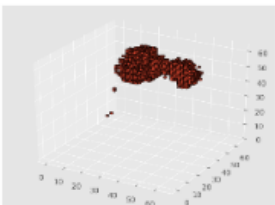
100回



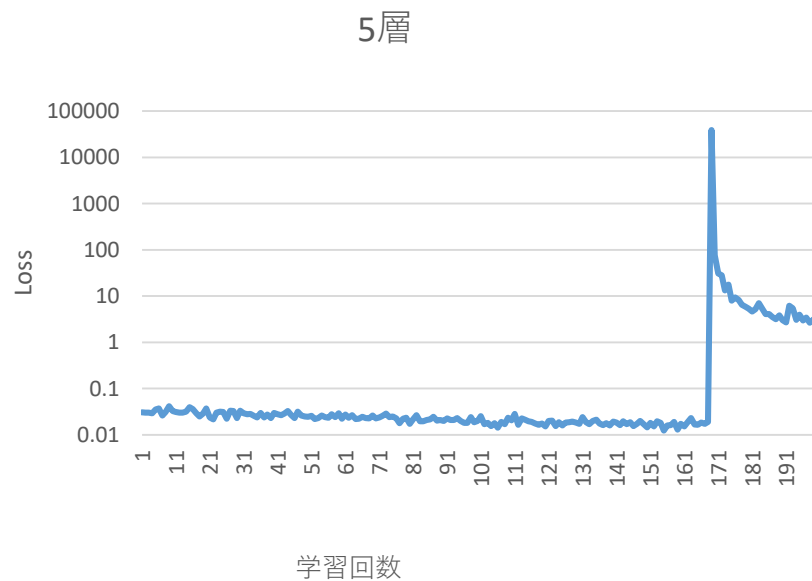
150回



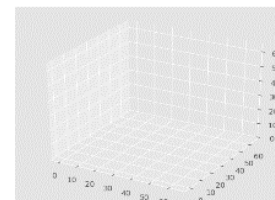
200回



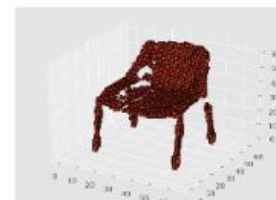
実験1：実験結果(5層)



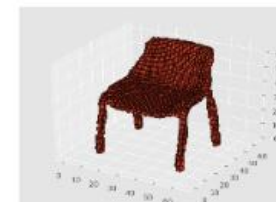
50回



100回



150回

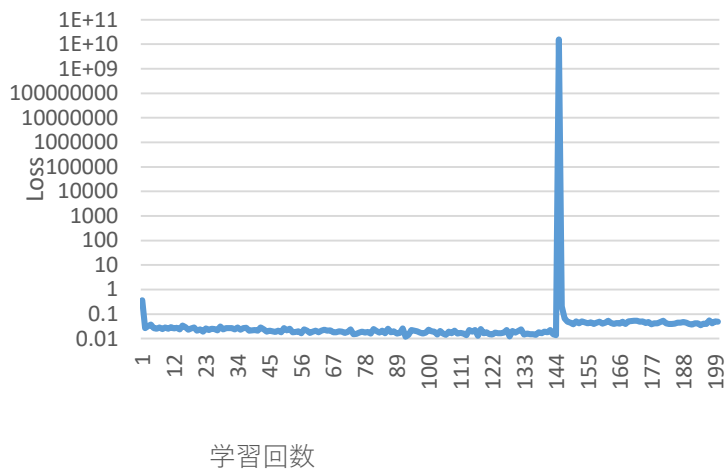


200回

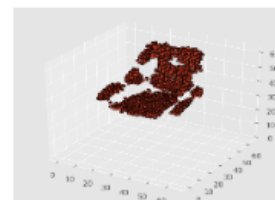


実験1：実験結果(6層)

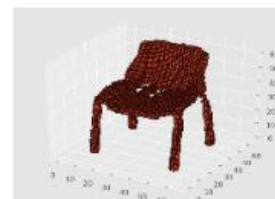
6層



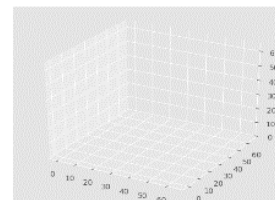
50回



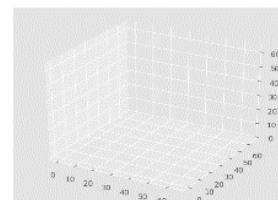
100回



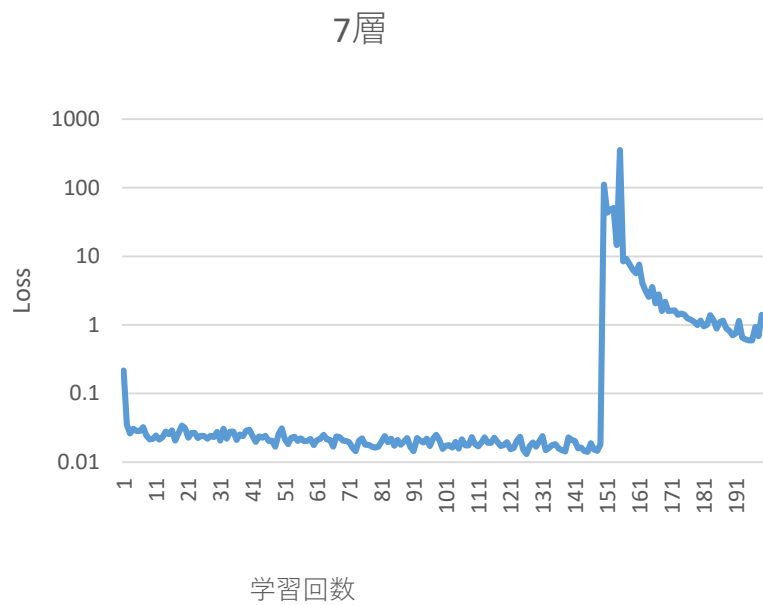
150回



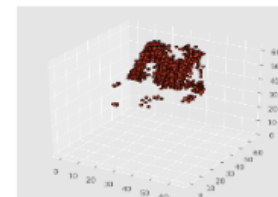
200回



実験1：実験結果(7層)



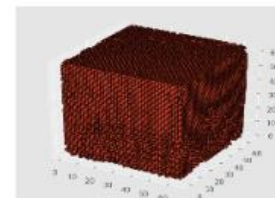
50回



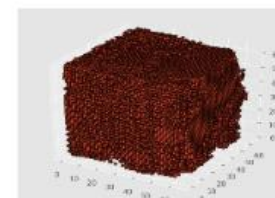
100回



150回



200回



実験1：実験結果(誤差)

生成結果と正解の同じ座標間でのボクセルの有無の違い

	3層	4層	5層	6層	7層
誤差の平均	8148	7776	7505	8331	7607
学習回数	50	100	150	100	100

層数は学習の安定化に影響しなかった
パラメータ数が多すぎるわけではなかった

実験2：実験方法

学習手法の変更により学習を安定化できるか調査した

		3D-SRGAN	3D-SRGAN_G1D5	3D-SRWGAN
変更	目的関数	×	×	○
	ハイパーパラメータ	×	○	○

実験2：実験方法

目的関数の変更

変更前

$$\mathbb{E}_{x \sim \mathbb{P}_r} [\log D(x)] + \mathbb{E}_{\tilde{x} \sim \mathbb{P}_g} [\log (1 - D(G(\tilde{x})))]$$



変更後

$$\mathbb{E}_{\tilde{x} \sim \mathbb{P}_g} [D(\tilde{x})] - \mathbb{E}_{x \sim \mathbb{P}_r} [D(x)] + \lambda \mathbb{E}_{\hat{x} \sim \mathbb{P}_{\hat{x}}} [(\|\nabla_{\hat{x}} D(\hat{x})\|_2 - 1)^2]$$

※WGAN-GPに合わせた

実験2：実験方法

ハイパーパラメータの変更

- 学習率

Generator : $2.5 \times 10^{-3} \rightarrow 2.0 \times 10^{-4}$

Discriminator : $1.0 \times 10^{-5} \rightarrow 2.0 \times 10^{-4}$

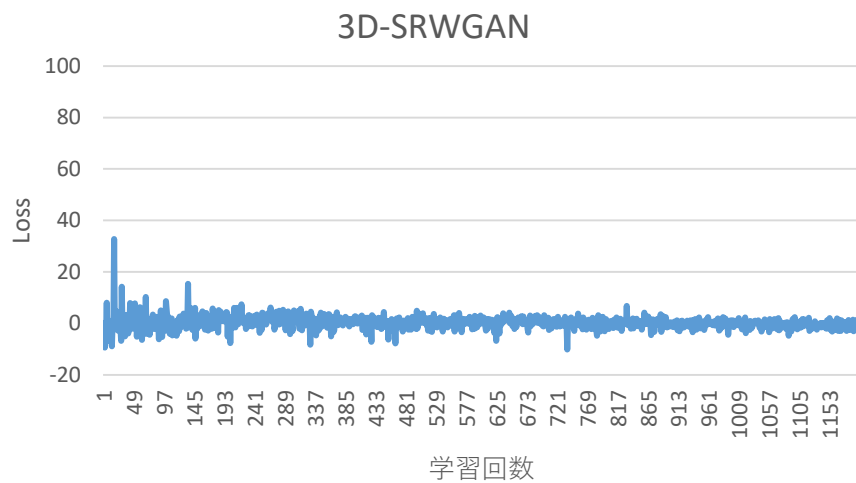
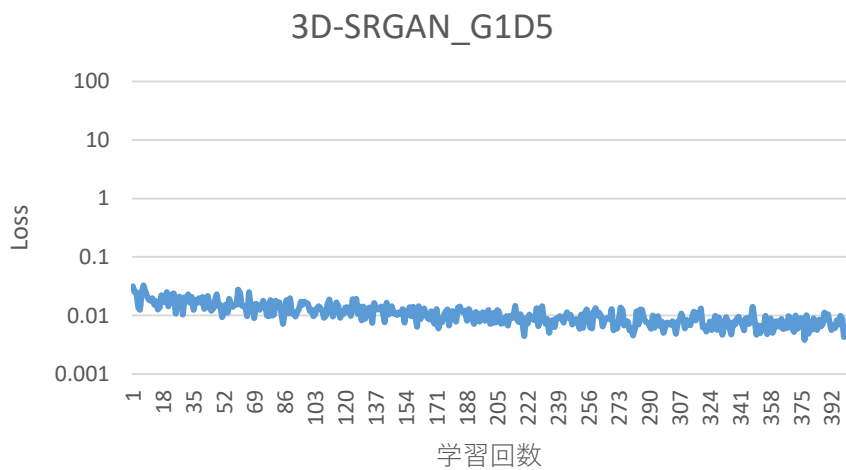
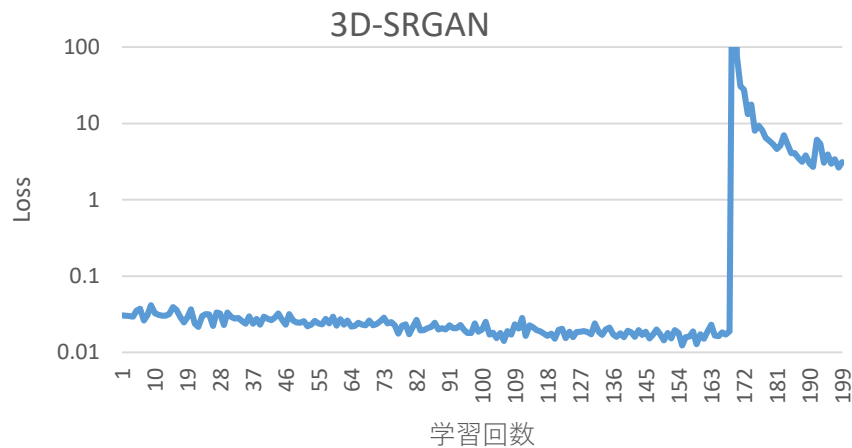
- GeneratorとDiscriminatorの更新頻度

1:1 \rightarrow 1:5

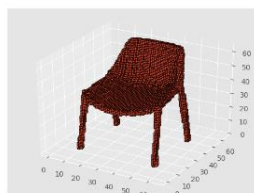
Generatorの学習を抑制

※WGAN-GPに合わせた

実験2の結果(Loss)



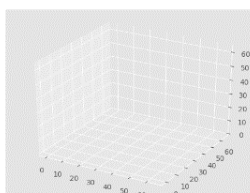
実験2の結果(生成結果)



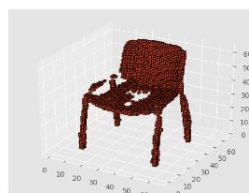
高解像3次元
モデル

50回

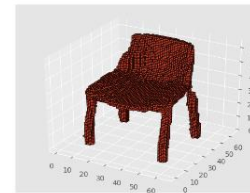
3D-SRGAN



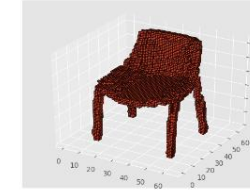
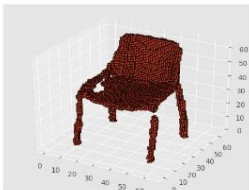
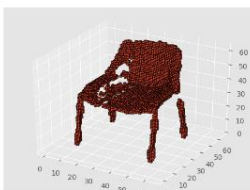
3D-SRGAN_
G1D5



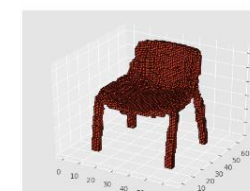
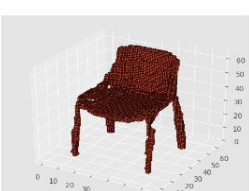
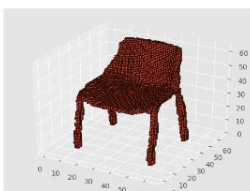
3D-SRWGAN



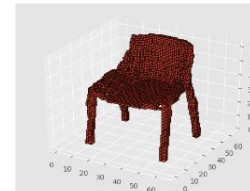
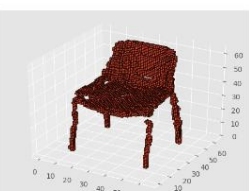
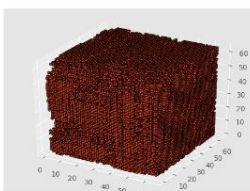
100回



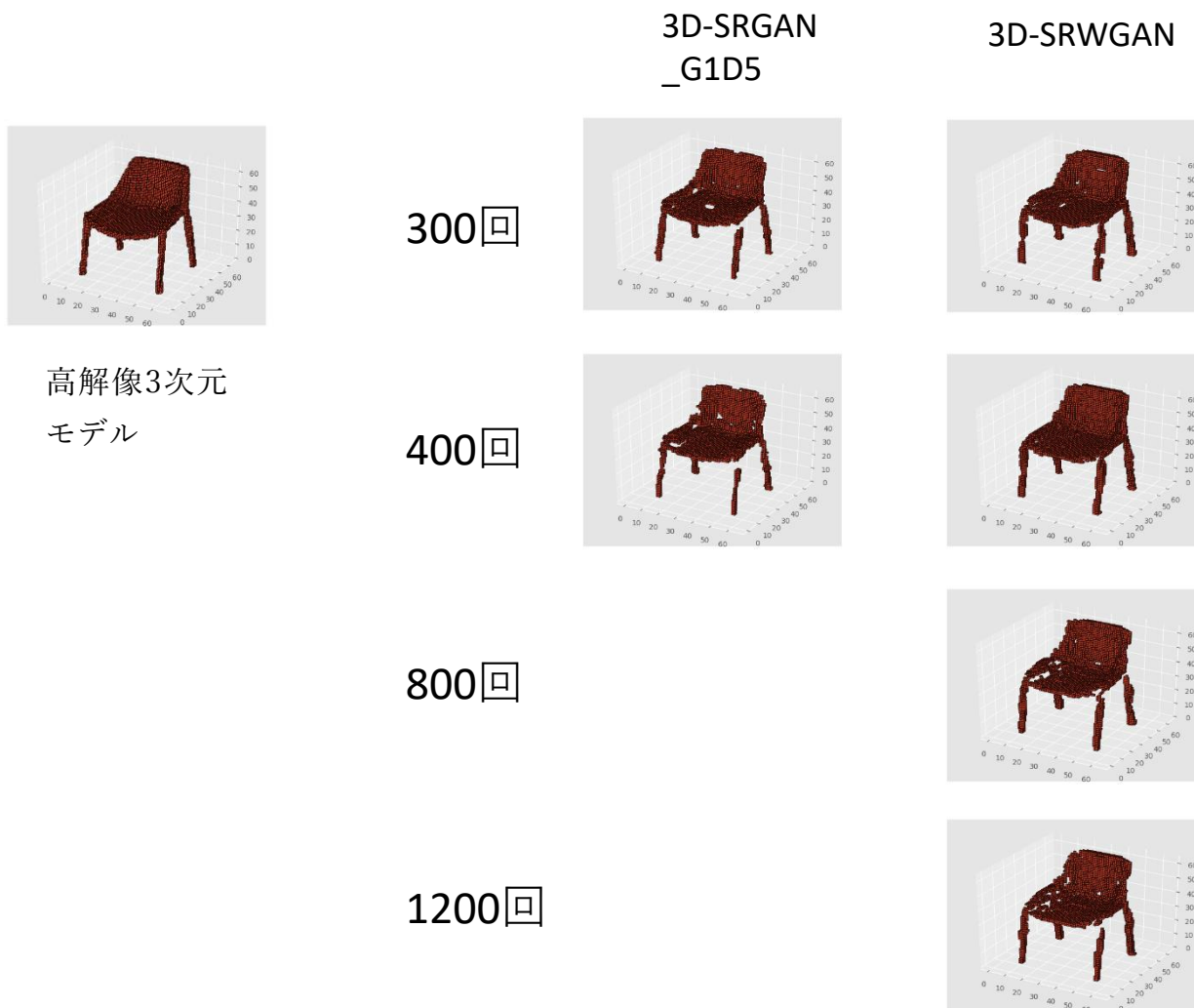
150回



200回



実験2：実験結果(生成結果)



実験2：実験結果(誤差)

生成結果と正解の同じ座標間でのボクセルの有無の違い

	3D-SRGAN	3D-SRGAN _G1D5	3D-SRWGAN
誤差の平均	7505	7045	7302
学習回数	150	250	1000

ハイパーパラメータの変更は学習の安定化に有効である

まとめ

3次元モデル超解像を行う3D-SRGANの学習を安定化させる方法の検討を行った

学習の安定化

- 層数の変更：影響しない
- ハイパーパラメータの変更：有効である
- 目的関数の変更：影響は確認できなかった

今後の課題

- ハイパーパラメータの細かい設定の見直し
- 目的関数の変更による影響について詳しく検討
- 安定化と精度の関係について詳しく検討