

クラス識別を導入した 3D-SRGANによるボクセル超解像

宮崎大学大学院 工学研究科 工学専攻

機械・情報系コース 情報システム工学分野

指導教員 椋木雅之

T2203425 野村淳也

2024/02/01

研究背景

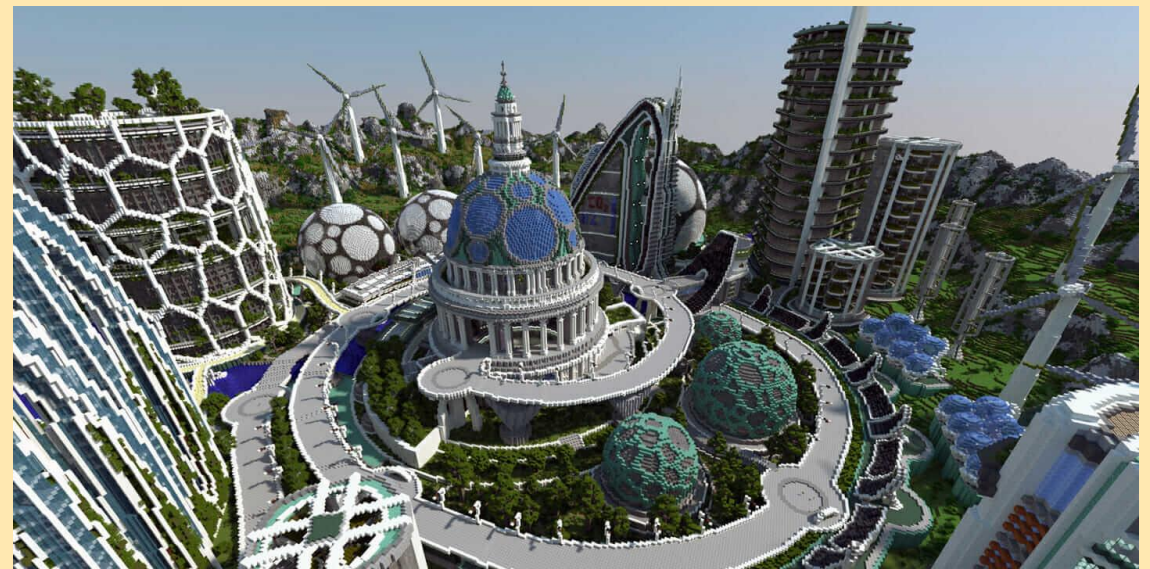
ボクセル表現の3次元モデリング

例)Minecraft

複雑なモデルの生成には
かなりの手間がかかる



引用: <https://mine-craftlife.com/entry/minecraft-tips-for-building>

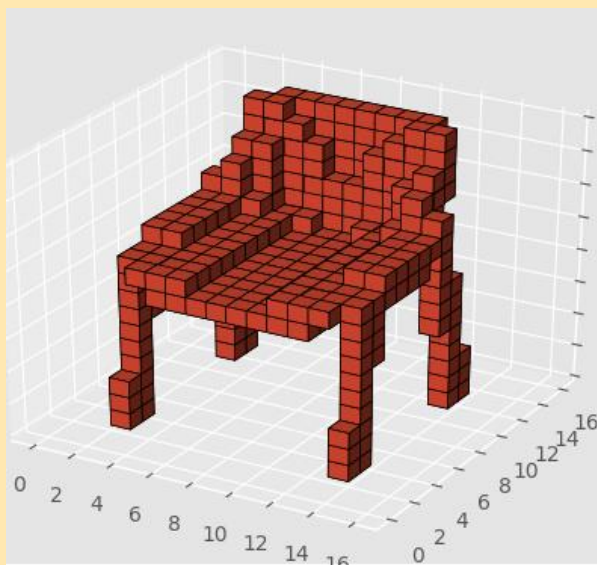


引用: <https://www.autodesk.co.jp/redshift/minecraft-architecture/>

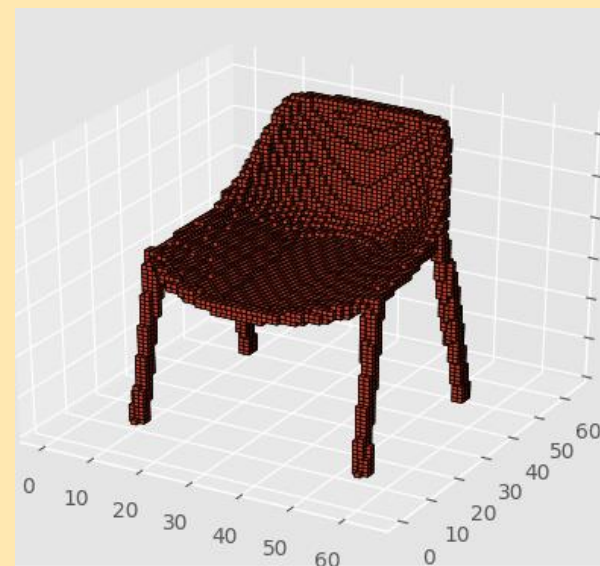
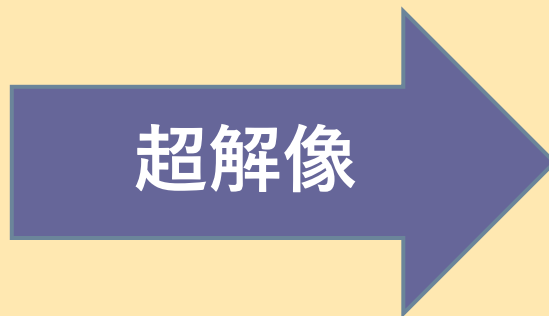
3次元ボクセルモデルの超解像

3D-SRGAN[1]

3次元ボクセルモデルの超解像を行う

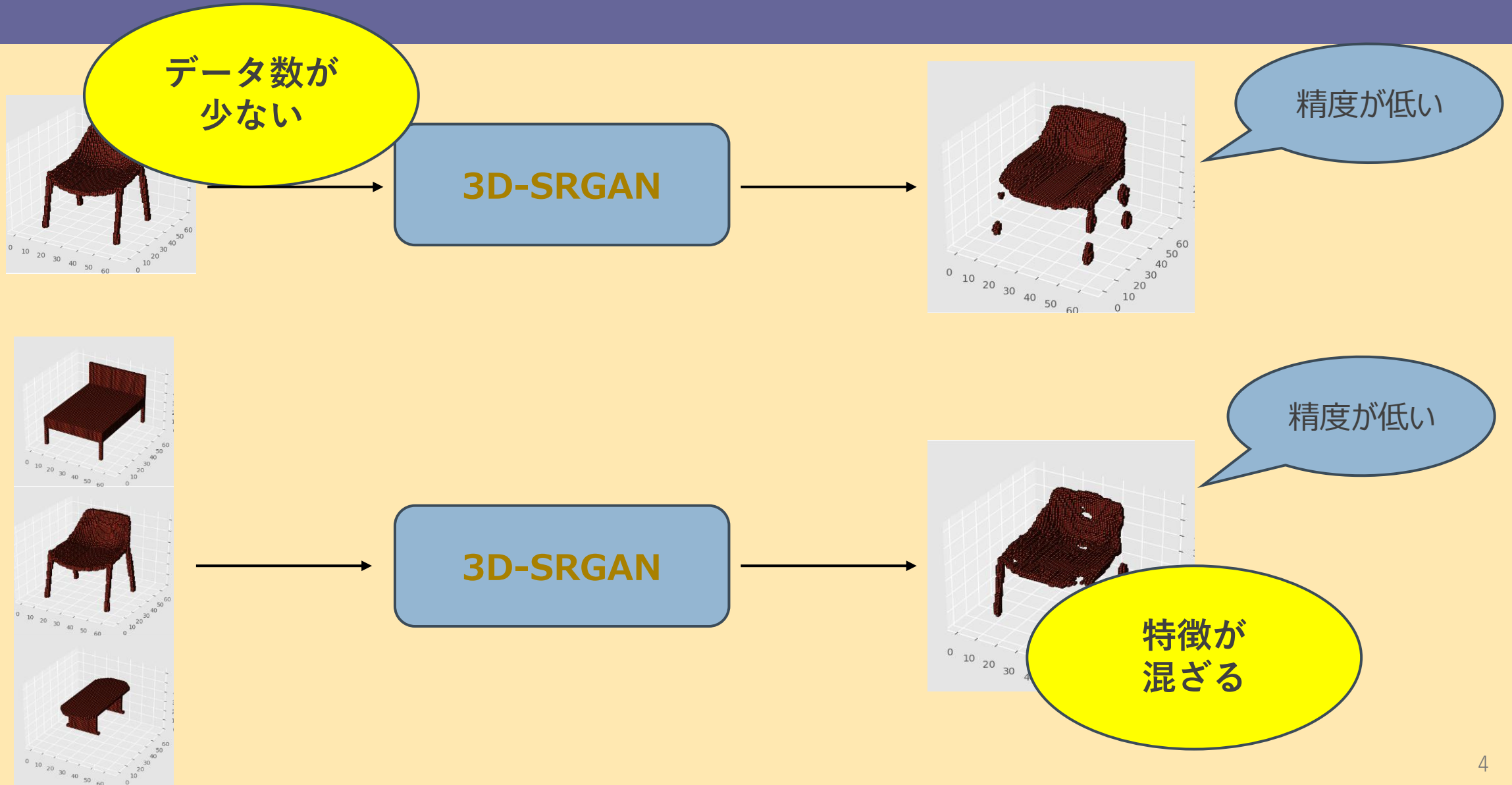


低解像度3次元モデル



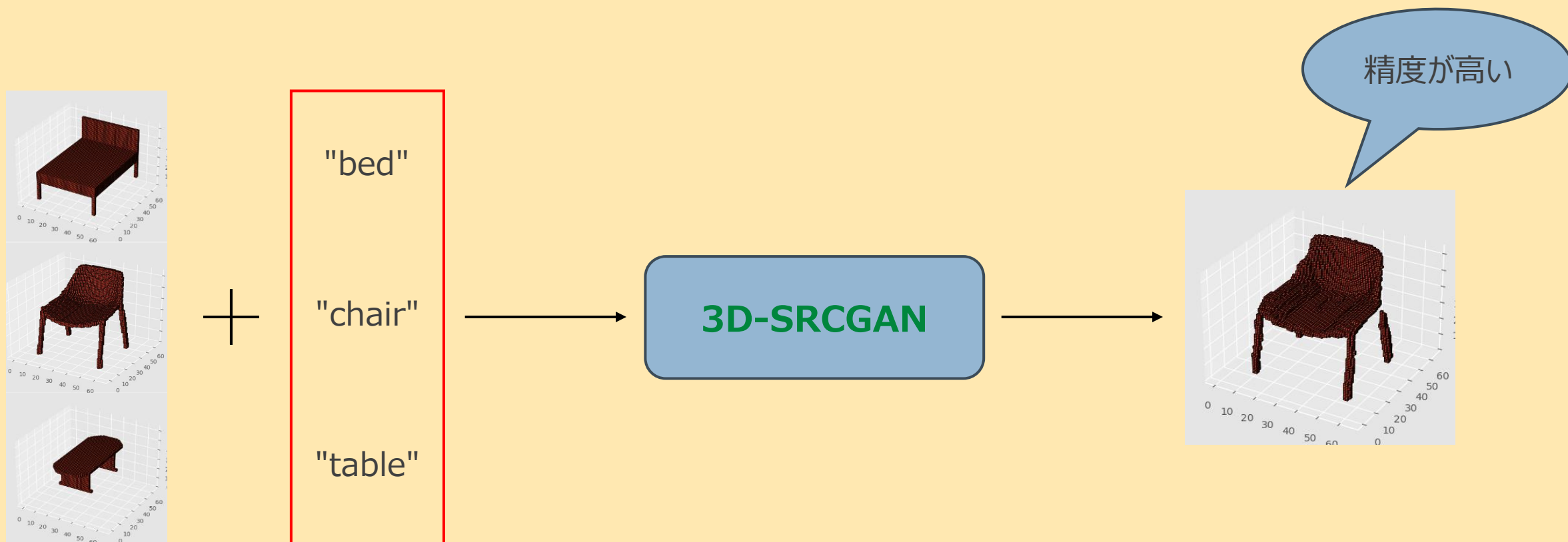
高解像度3次元モデル

従来手法と問題点

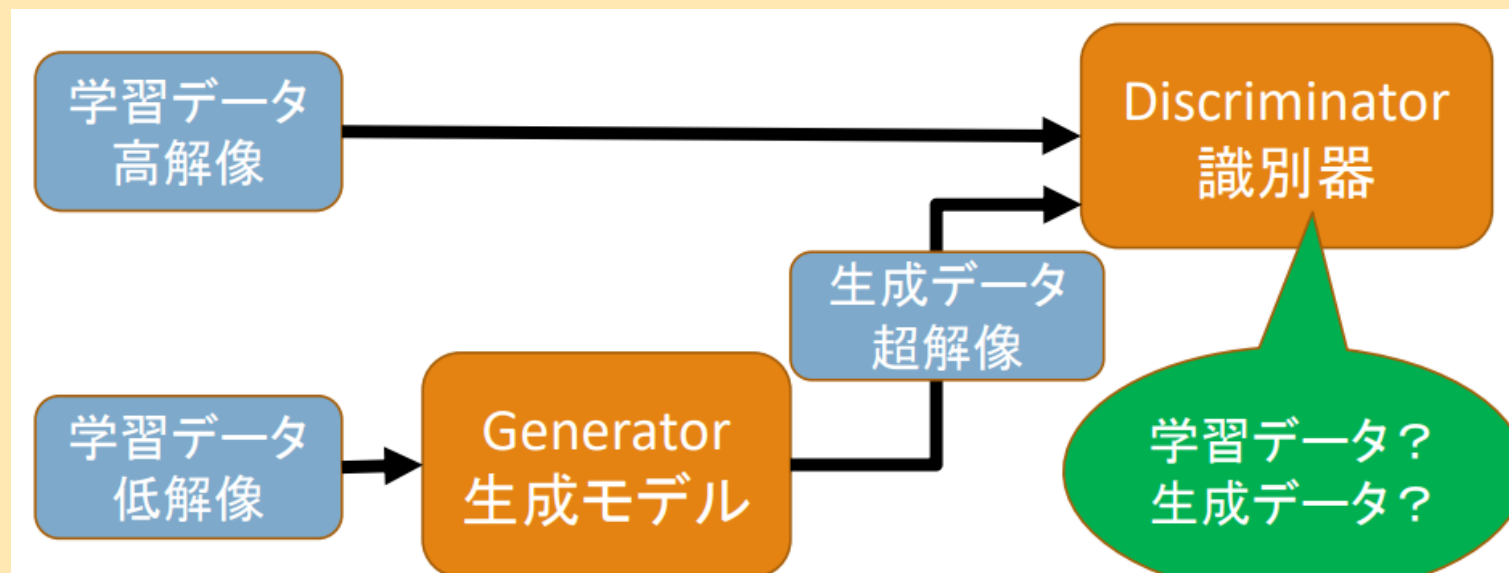


目的

3D-SRGANにクラス識別情報を導入し、超解像精度を改善する



3D-SRGAN

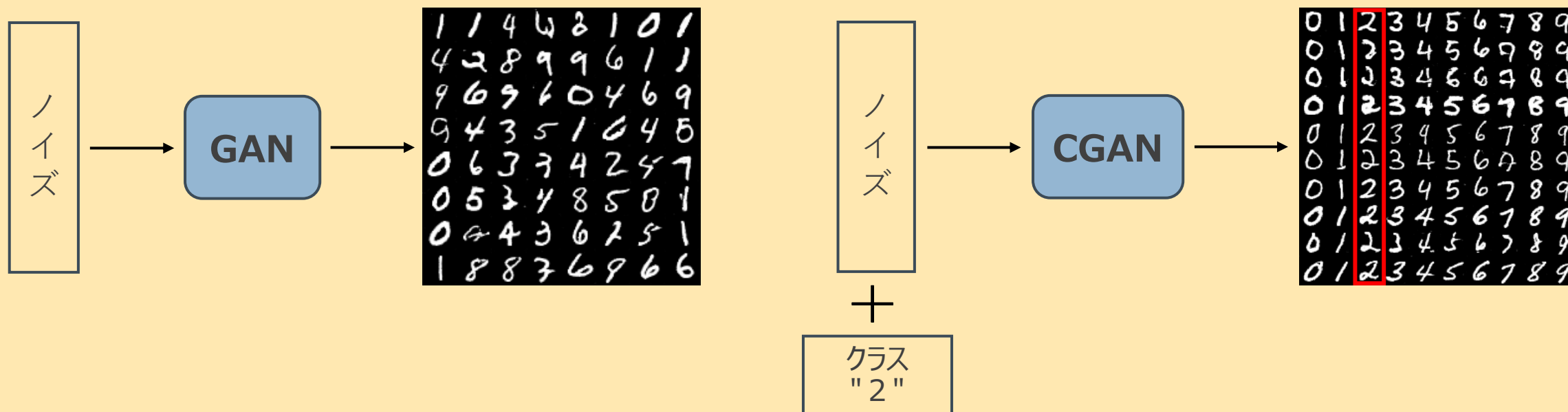


- 2次元画像の超解像を行うSRGAN[2]を3次元に拡張
- Generatorが学習データに近いモデルを生成することが目的

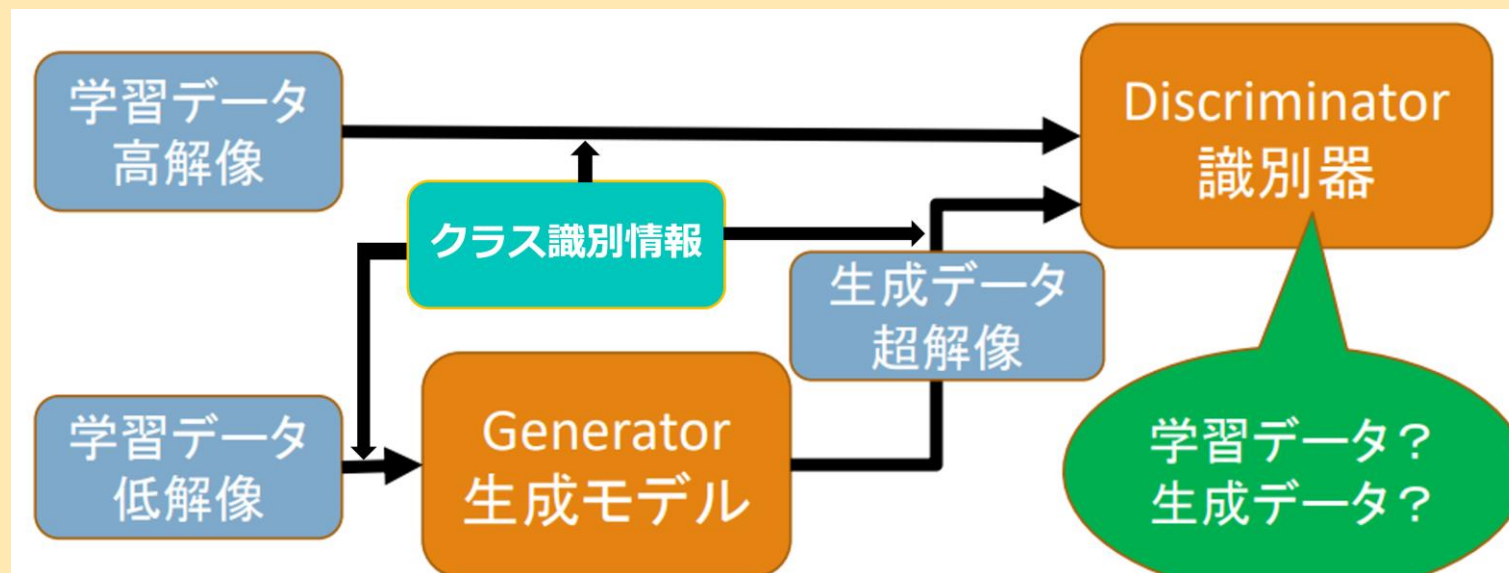
従来手法

CGAN

- Mirza[3]らによって提案された手法
- クラス識別情報によって生成画像の制御を行う



3D-SRCGAN



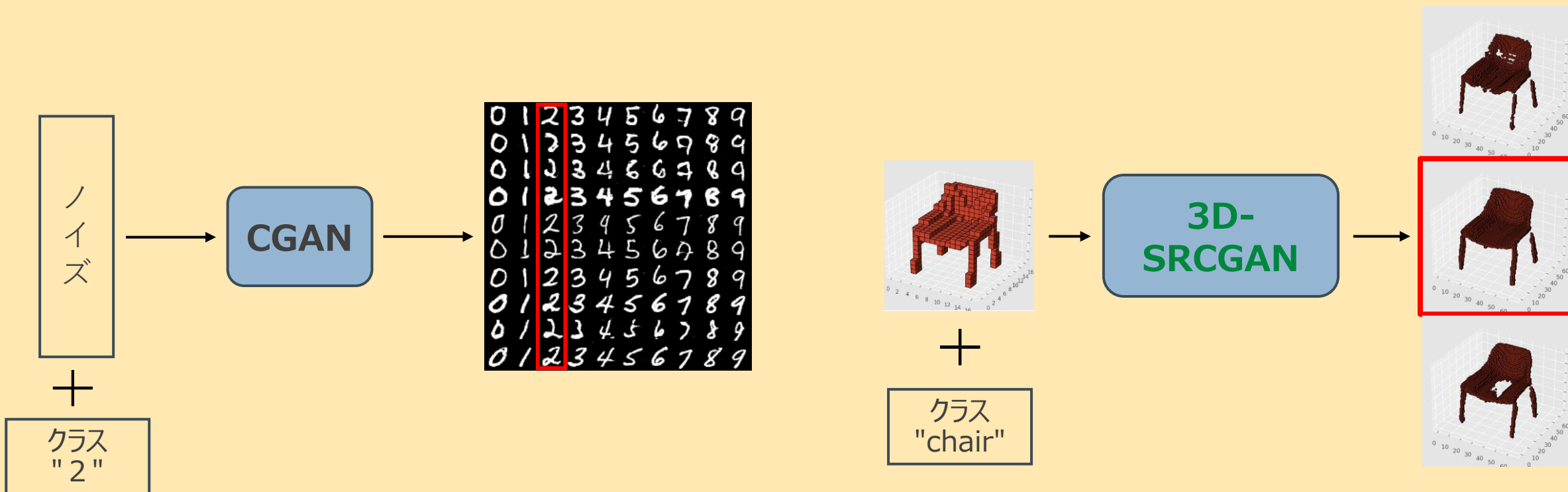
- GeneratorとDiscriminatorの入力にクラス識別情報を付加

提案手法

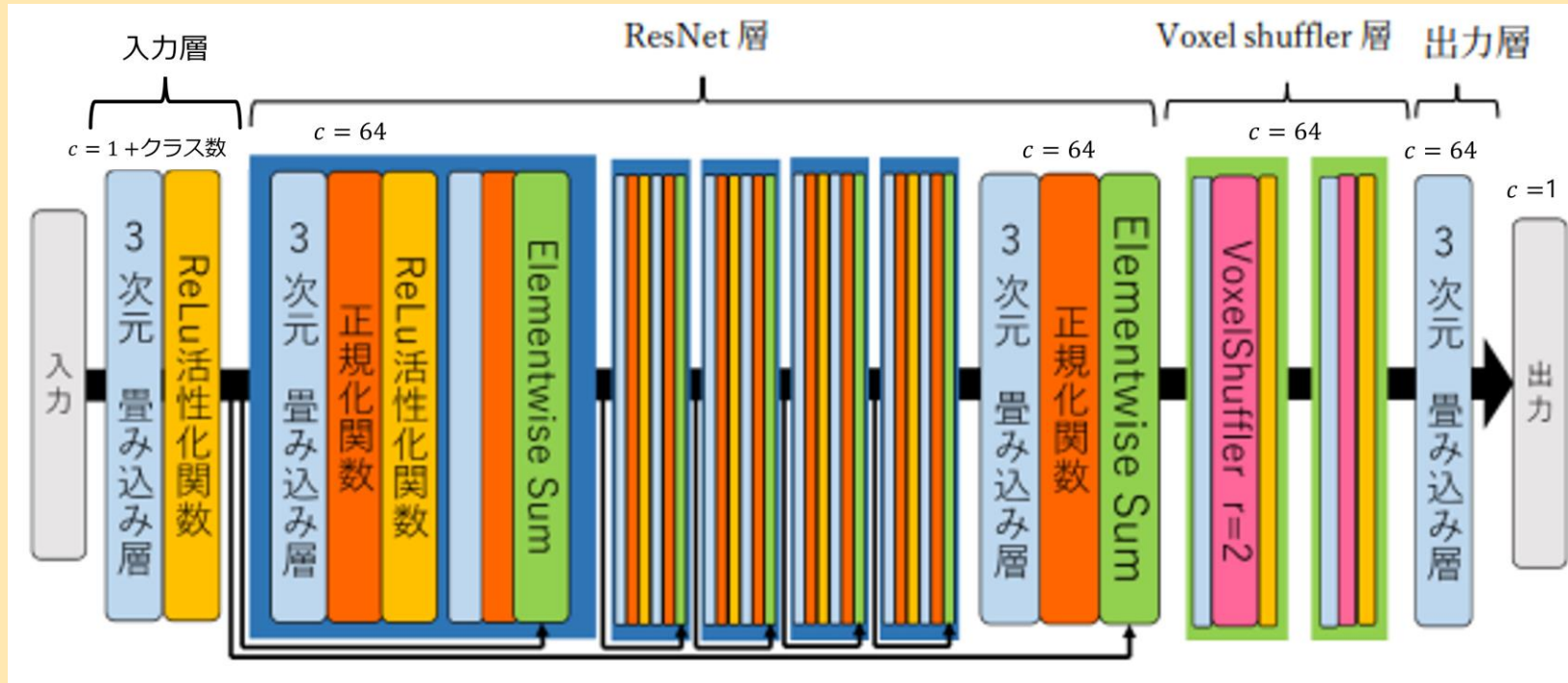
3D-Super Resolution Conditional Generative Adversarial Networks

3D-SRCGAN

- クラス識別情報によって各クラスの特徴が混ざらない

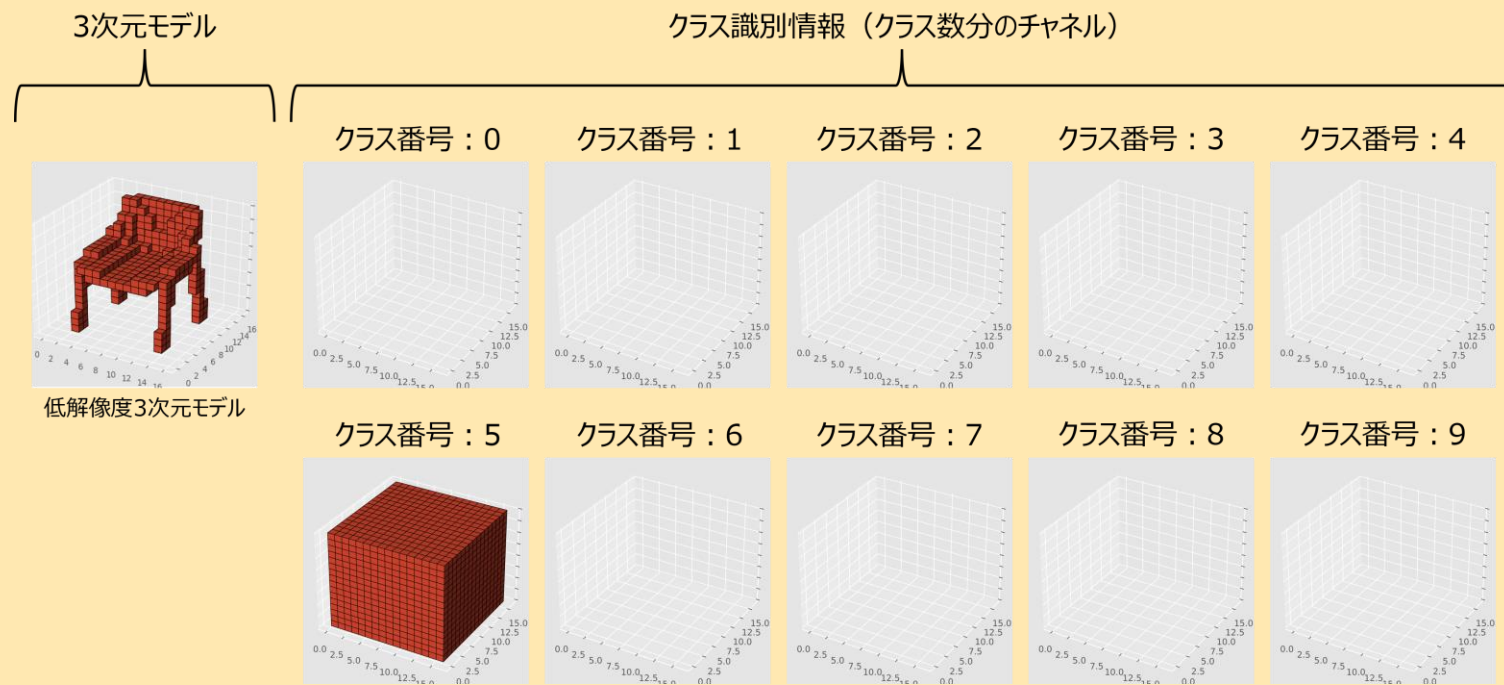


Generatorのネットワーク構造



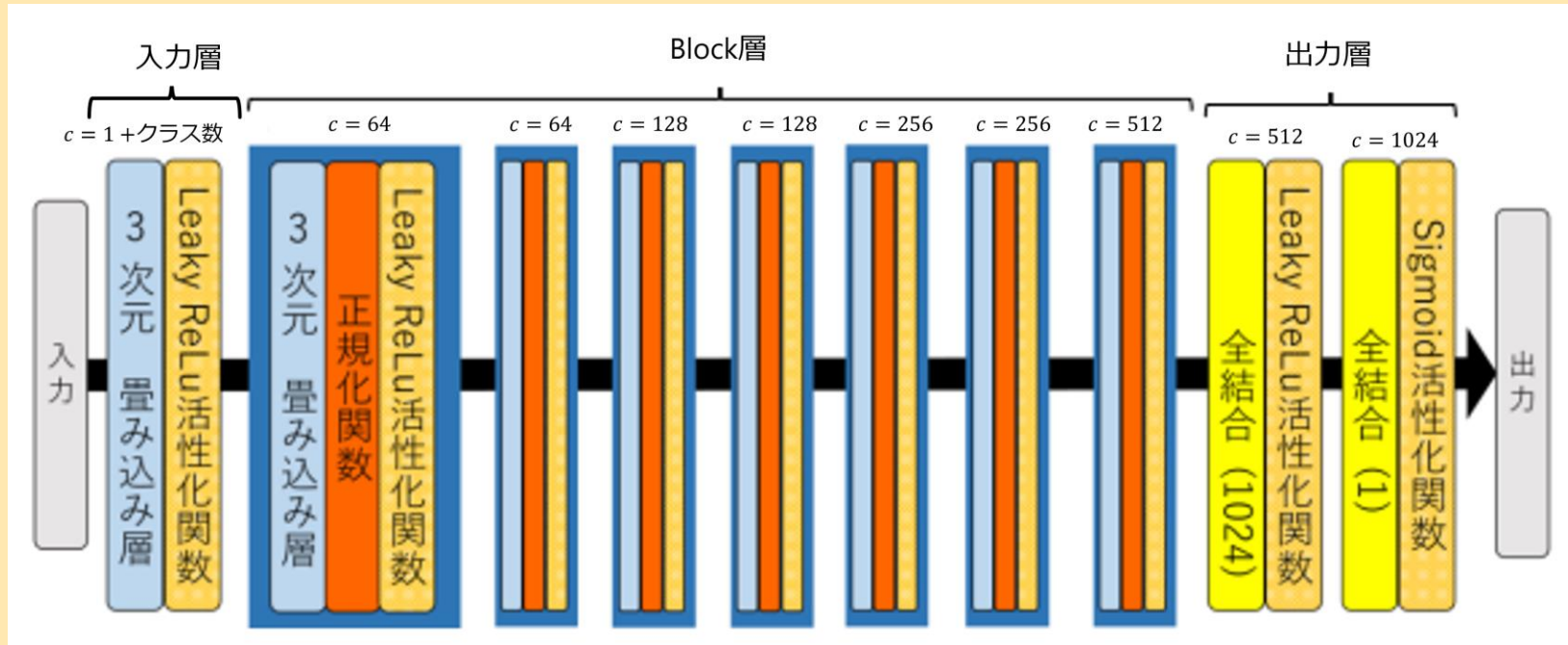
- Generatorの構造は3D-SRGANと同様
- 入力層への入力にクラス識別情報を付加する

クラス識別情報の付加



- Generatorの入力（低解像度3次元モデル）に付加
- クラスによってボクセルで埋める場所を変える

Discriminatorのネットワーク構造



- Discriminatorの構造は3D-SRGANと同様
- 入力層への入力にクラス識別情報を付加する

クラス識別情報の付加



- Discriminatorの入力 (超解像3次元モデル) に付加
- クラスによってボクセルで埋める場所を変える

実験

実験 1 クラス識別情報の導入による精度向上の評価

実験 2 複数クラスで学習することによる精度向上の評価

実験 3 データセットの解像度を上げた場合の影響調査

評価方法

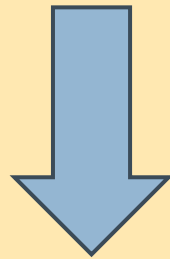
高解像度 3 次元モデルと超解像 3 次元モデルの

- 連結成分の数の誤差 (連)
- 穴の数の誤差 (穴)
- 空洞の数の誤差 (空)
- 同じ座標間のボクセルの値の誤差 (値)

実験データ

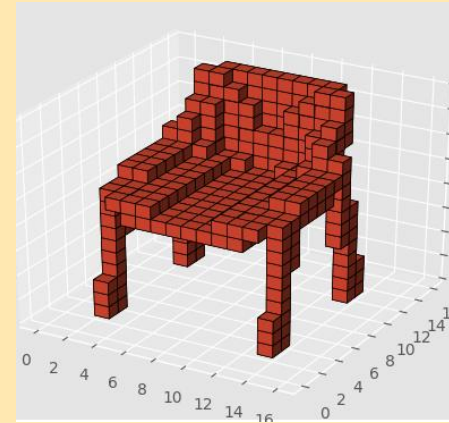
ModelNet10[4]の3次元モデル

- ・ 学習データ : 3991個
- ・ テストデータ : 908個

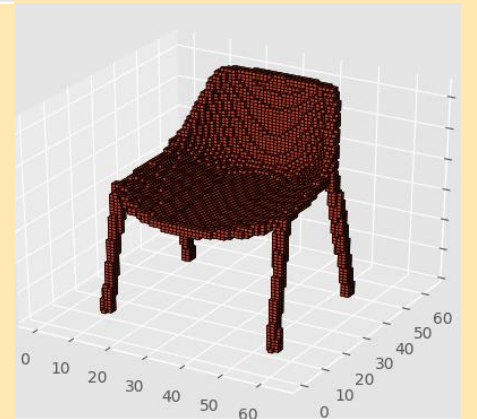


低解像度3次元モデル : $16 \times 16 \times 16$

高解像度3次元モデル : $64 \times 64 \times 64$



低解像度3次元モデル

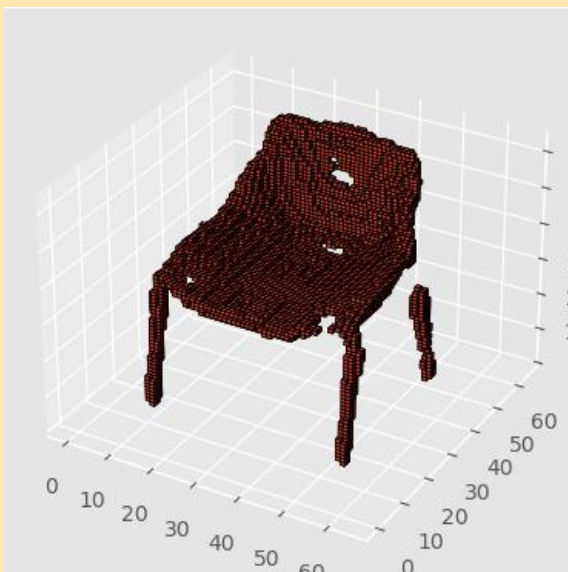


高解像度3次元モデル

実験1 – 実験内容

- クラス識別情報の導入による精度向上の評価
- 比較対象
 - クラス識別情報がない従来の**3D-SRGAN**
 - クラス識別情報を導入した**3D-SRCGAN**
 - ModelNet10の10クラスすべてのデータを使用

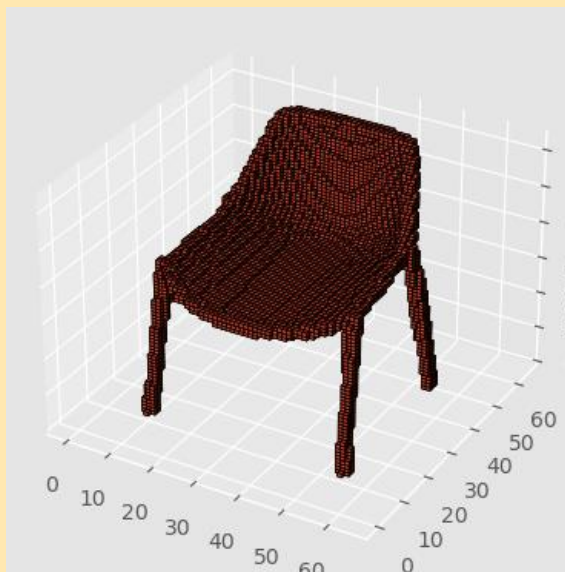
実験1 - 超解像結果



3D-SRGAN

定量評価

連 : 4
穴 : 21
空 : 0
値 : 11587



高解像度3次元モデル

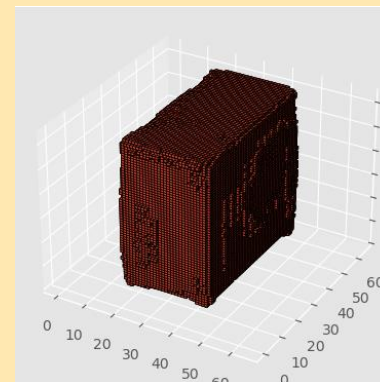
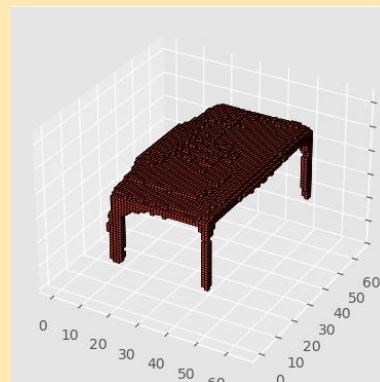
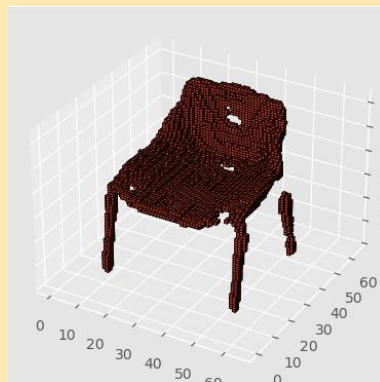
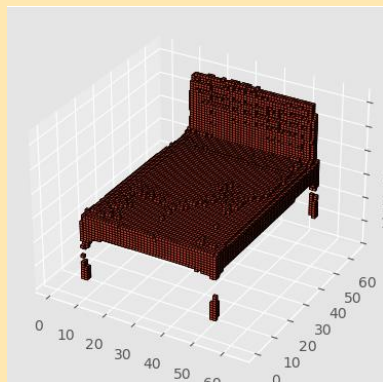
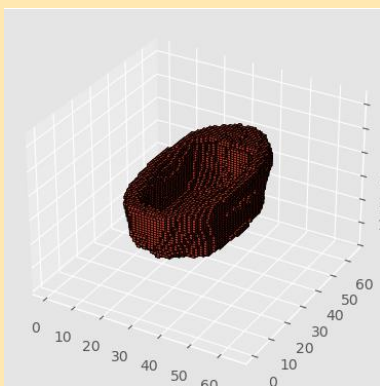
3D-SRCGAN

定量評価

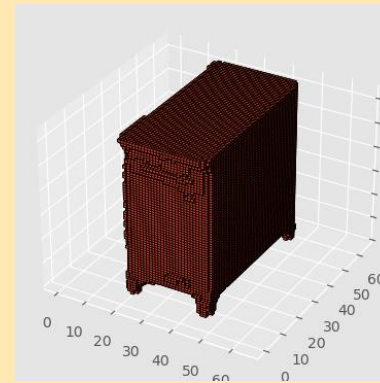
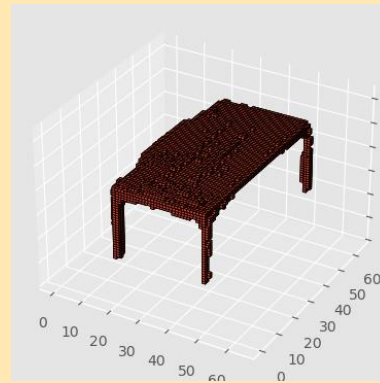
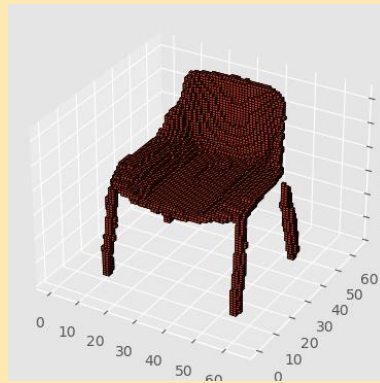
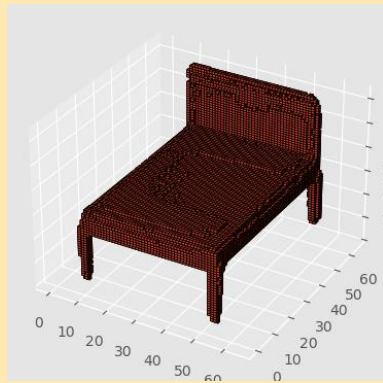
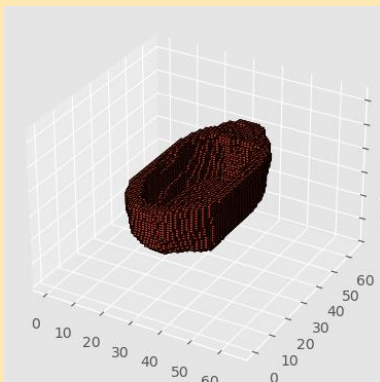
連 : 3
穴 : 12
空 : 0
値 : 2982

実験1 - 超解像結果

3D-SRGAN

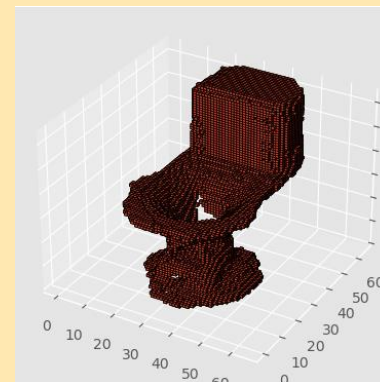
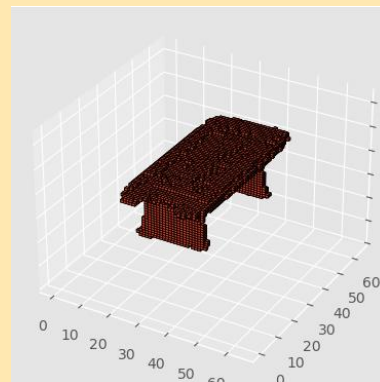
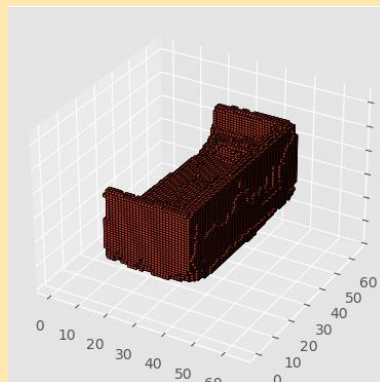
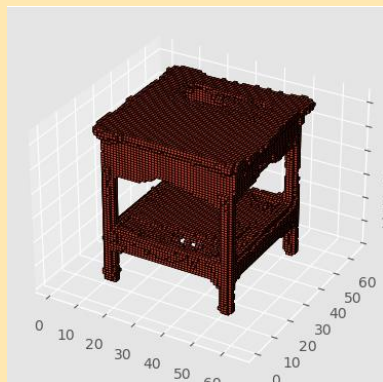
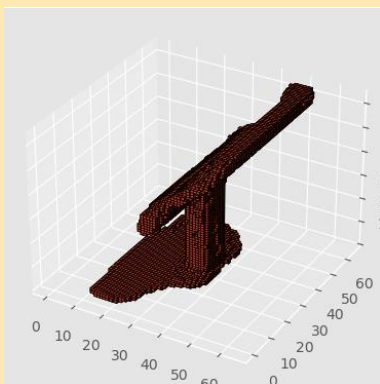


3D-SRCGAN

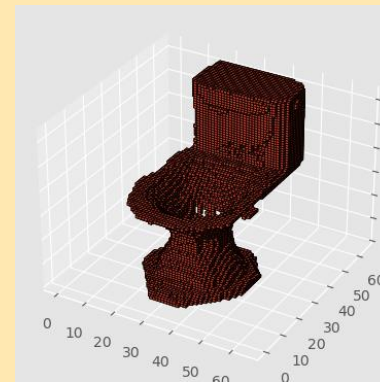
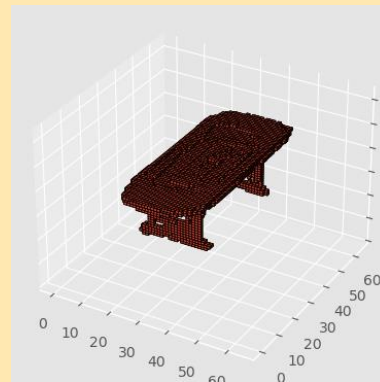
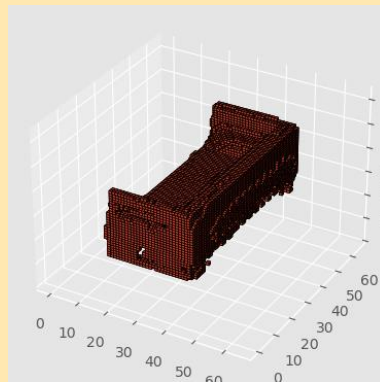
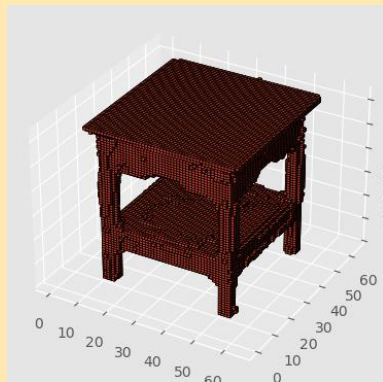
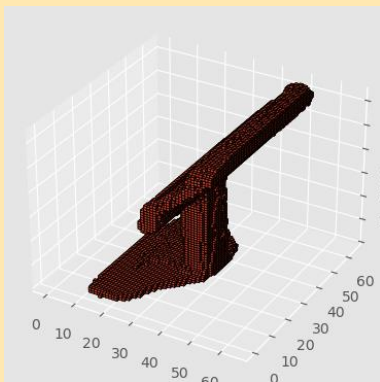


実験1 - 超解像結果

3D-SRGAN



3D-SRCGAN



実験1 – 定量評価

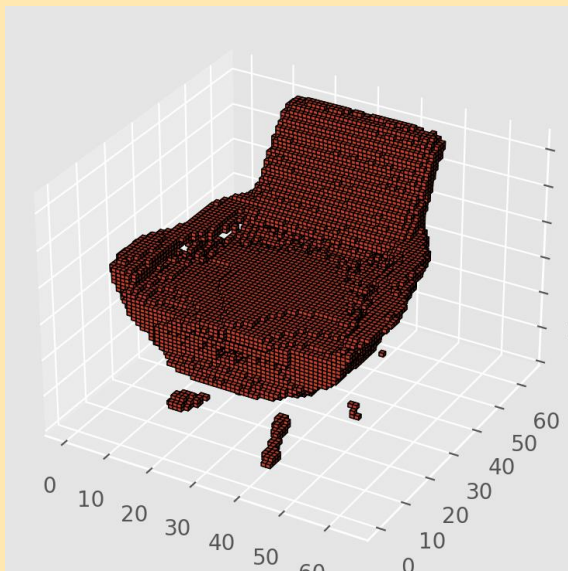
	連	穴	空	値
3D-SRGAN	2.7	6.9	1.8	47640.2
3D-SRCGAN	2.2	5.6	1.3	10189.3

- 10クラスのテストデータ全体に対する平均値
- 全ての項目で**3D-SRCGAN**が優れている
 - ボクセルの値は特に大きな差がある

実験2 – 実験内容

- 複数クラスで学習することによる精度向上の評価
- 比較対象
 - クラス識別情報がない従来の**3D-SRGAN**
 - ModelNet10の**chairのクラスのみ**のデータを使用
 - クラス識別情報を導入した**3D-SRCGAN**
 - ModelNet10の**10クラスすべて**のデータを使用

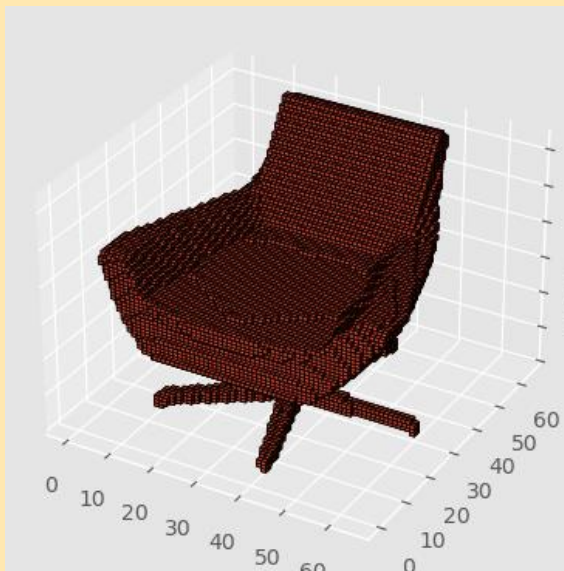
実験2 - 超解像結果



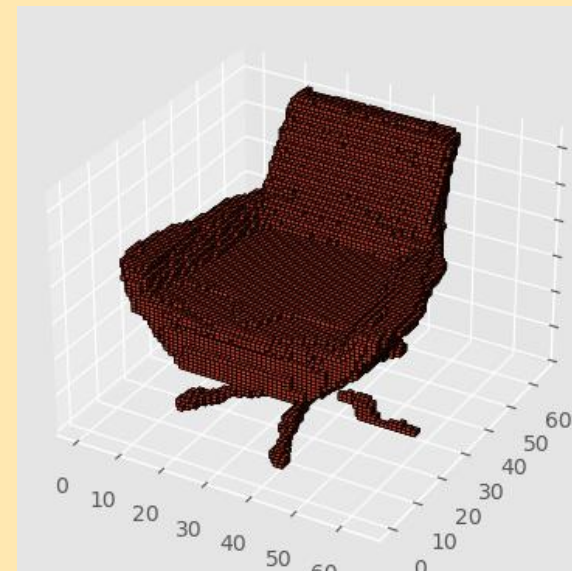
3D-SRGAN

定量評価

連 : 51
穴 : 31
空 : 7
値 : 24112



高解像度3次元モデル



3D-SRCGAN

定量評価

連 : 2
穴 : 11
空 : 0
値 : 6087

実験2 – 定量評価

	連	穴	空	値
3D-SRGAN (1クラス)	5.9	8.5	0.7	7893.8
3D-SRCGAN (10クラス)	2.6	4.4	0.2	5968.6

- 椅子クラスのテストデータ全体に対する平均値
- 全ての項目で**3D-SRCGAN**が優れている
 - 連結成分、穴、空洞の誤差の差が大きい

実験3 – 実験内容

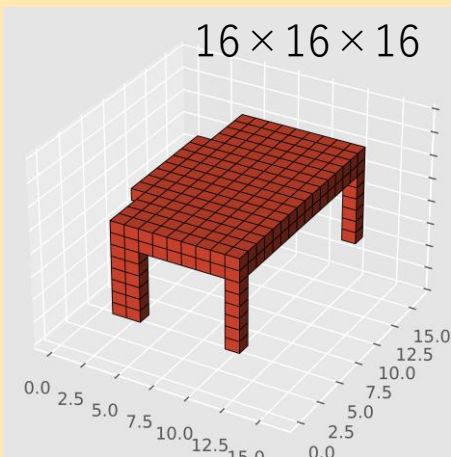
- データセットの解像度を上げた場合の影響調査
- 比較対象
 - 従来 of 解像度 ($16^3 \rightarrow 64^3$) で学習した **3D-SRCGAN**
 - 変更後の解像度 ($32^3 \rightarrow 128^3$) で学習した **3D-SRCGAN**
 - desk、monitor、night_stand、table のクラスのデータを使用

実験3 - 超解像結果

従来の解像度

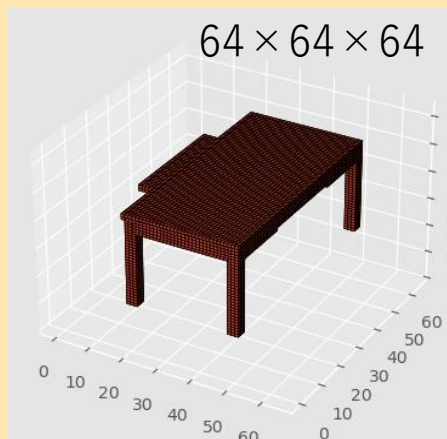
低解像度3次元モデル

$16 \times 16 \times 16$



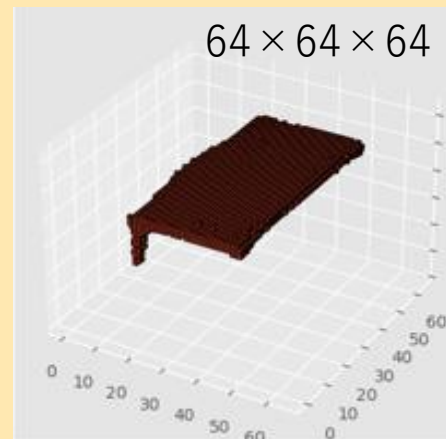
高解像度3次元モデル

$64 \times 64 \times 64$



超解像3次元モデル

$64 \times 64 \times 64$

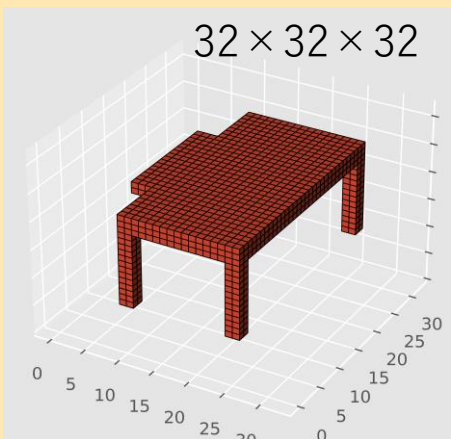


定量評価

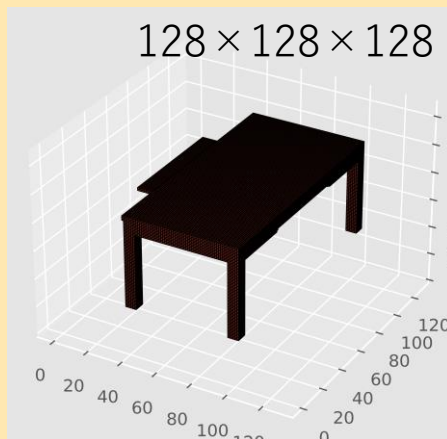
連 : 1
穴 : 0
空 : 0
値 : 50984

変更後

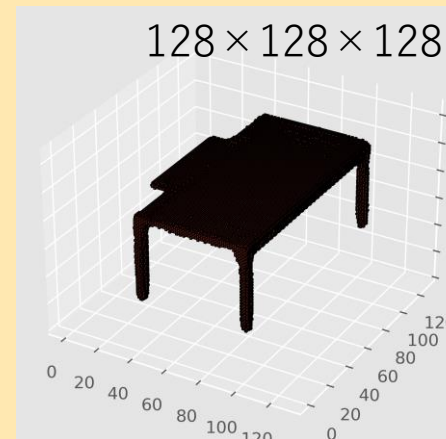
$32 \times 32 \times 32$



$128 \times 128 \times 128$



$128 \times 128 \times 128$

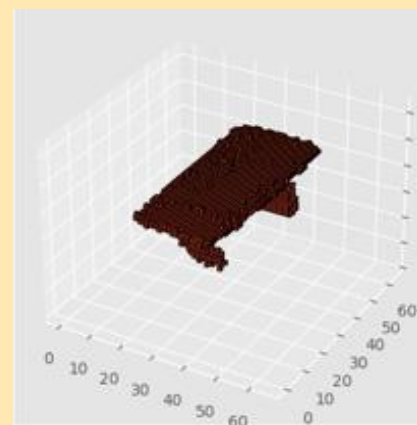
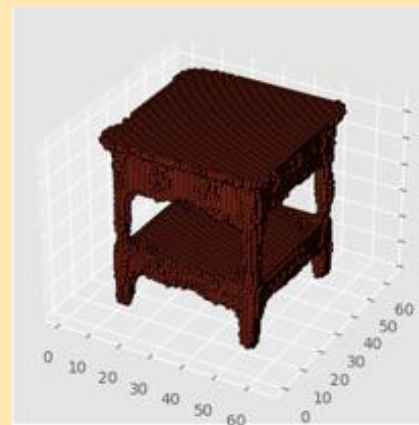
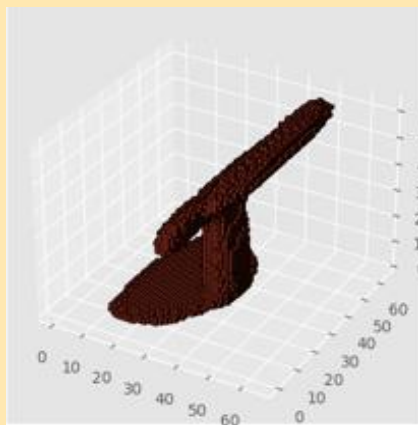
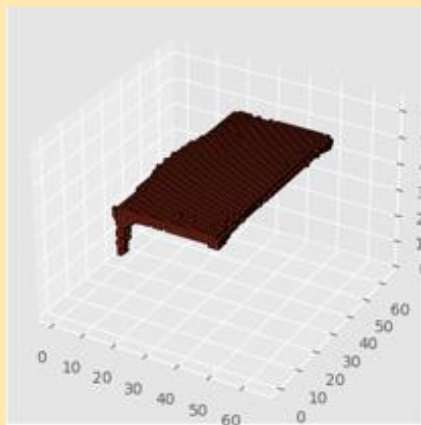


定量評価

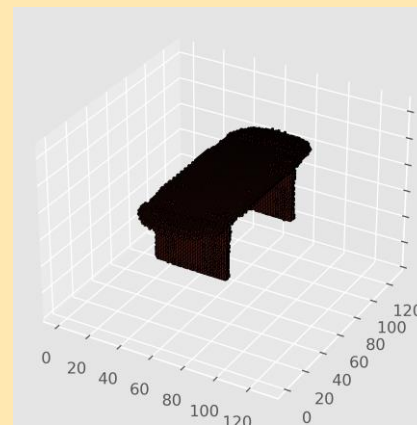
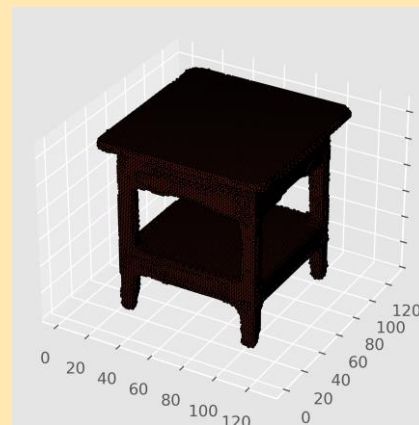
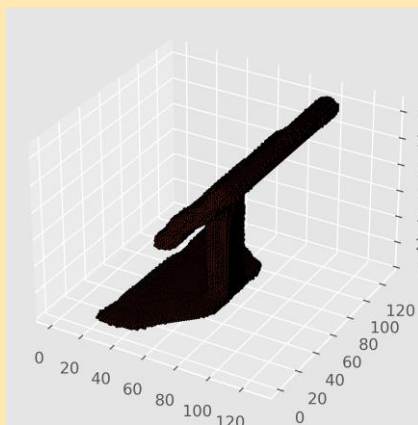
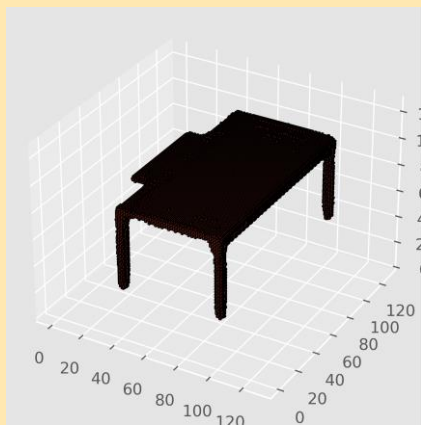
連 : 10
穴 : 6
空 : 0
値 : 20397

実験3 - 超解像結果

従来
の解像度
(64×64×64)



変更後
(128×128×128)



実験3 – 定量評価

	連	穴	空	値
従来 of 解像度 (64×64×64)	143.2	157.6	49.6	90023.2
変更後 (128×128×128)	95.4	99	21.7	55872.2

- 4クラスのテストデータ全体に対する平均値
- 64×64×64は値のみ8倍している
- 全ての項目で解像度を上げた場合が優れている

まとめ

従来の**3D-SRGAN**にクラス識別を導入した**3D-SRCGAN**を提案した

- 超解像精度を落とさず、学習データ数を増やせる
- データ数の増加により、超解像精度が向上する
- 複数クラスの3次元モデルを高精度に超解像することができる
- データセットの解像度を上げることで超解像精度、品質ともに向上する

今後の課題

- クラス識別情報の付加方法を変更する
- 3次元モデリング以外の分野に応用する

参考文献

- [1]岡和寿, 椋木雅之, “SRGAN の 3 次元ボクセルモデル超解像への適用”, 画像電子学会論文誌, vol. 48, no. 4, pp. 448-496, (2019)
- [2] C Ledig, L Theis, F Huszar, J Caballero, A Cunningham, A Acosta, "Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network", IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 105-114, (2017)
- [3] M Mirza, S Osindero, “Conditional generative adversarial nets” arXiv:1411.1784 (2014)
- [4] ModelNet10, <http://modelnet.cs.princeton.edu>