

動画像からの物体検出のための U-Net3Dの改良

2022/2/16 卒業論文発表会
宮崎大学工学部 情報システム工学科
指導教員 椋木雅之
67180320 中山隼人

研究背景

- ・ 物体検出の需要が増えている。
- ・ 深層学習を用いた物体検出は精度がかなり高い。

Ex) ・ マーケティング用途への応用

→ カメラの映像から顧客がどのような商品に興味を持っていて、実際にどんな商品を買ったか。

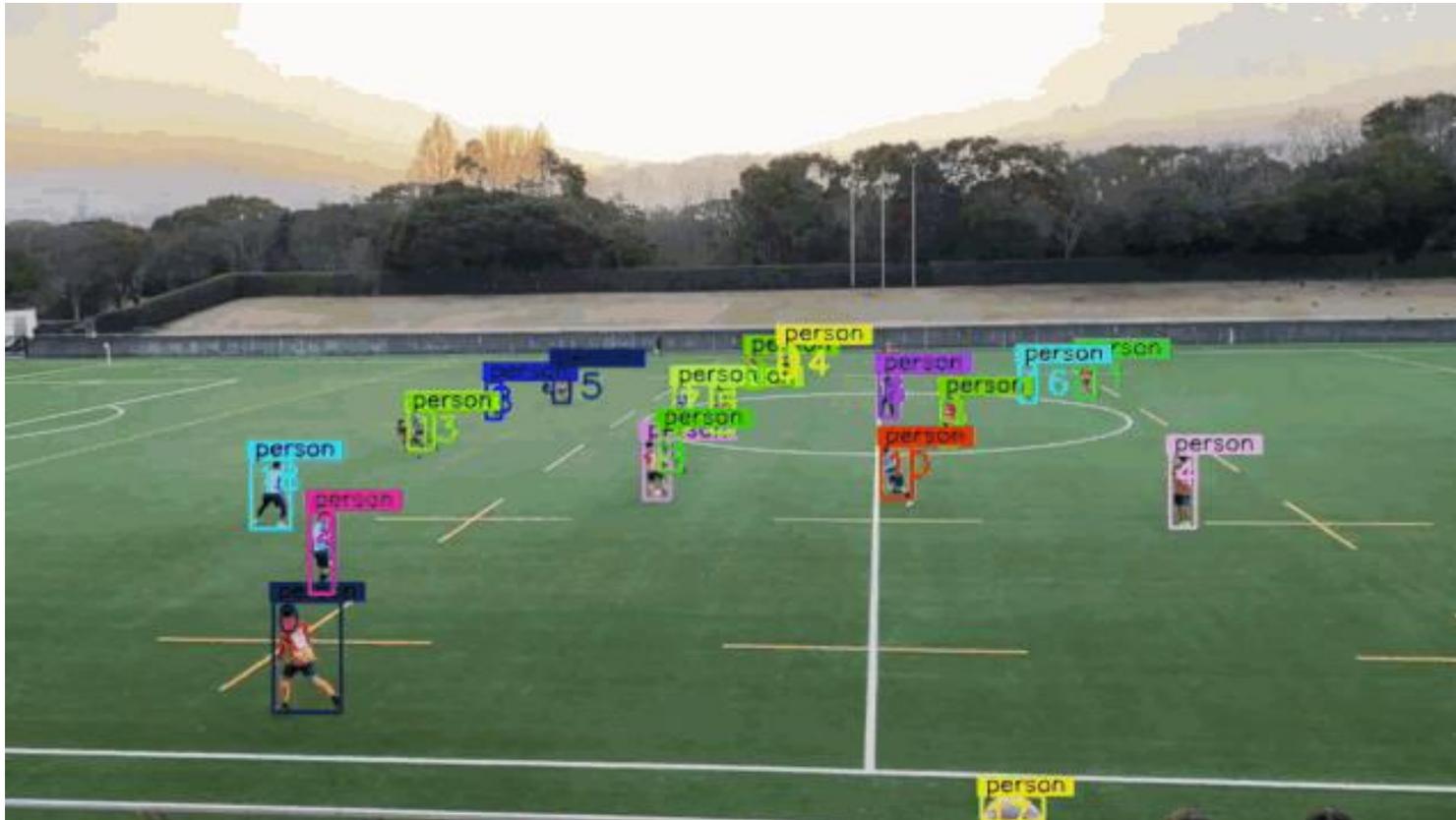
- ・ 自動運転
- ・ 農業への活用

➡ **しかし、連続検出の精度が悪い**



従来研究①

- 動画像からの物体検出は途中で検出漏れが出てくる[11]



※イメージ：パラパラ漫画

研究目的

時間方向での連続検出の精度を上げたい



時間方向に束ねて処理（3次元画像で）



しかし、従来の**U-Net**では良くならない



U-Netを改良する



U-Netとは

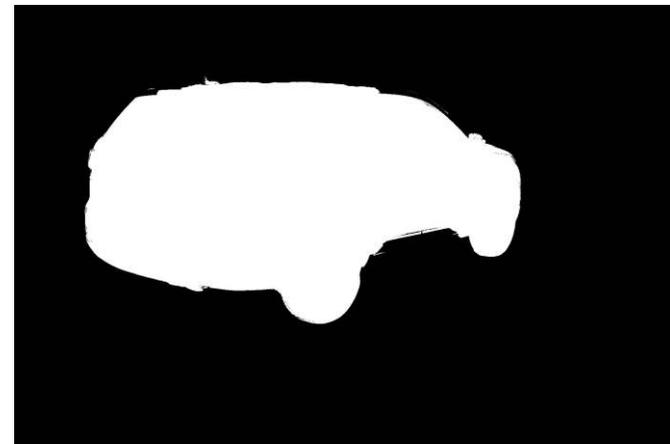
- 物体検出をするネットワークの1つ

→ **ネットワーク構造が単純**

→ **他のネットワークよりも学習データが少なくても、精度が高い**

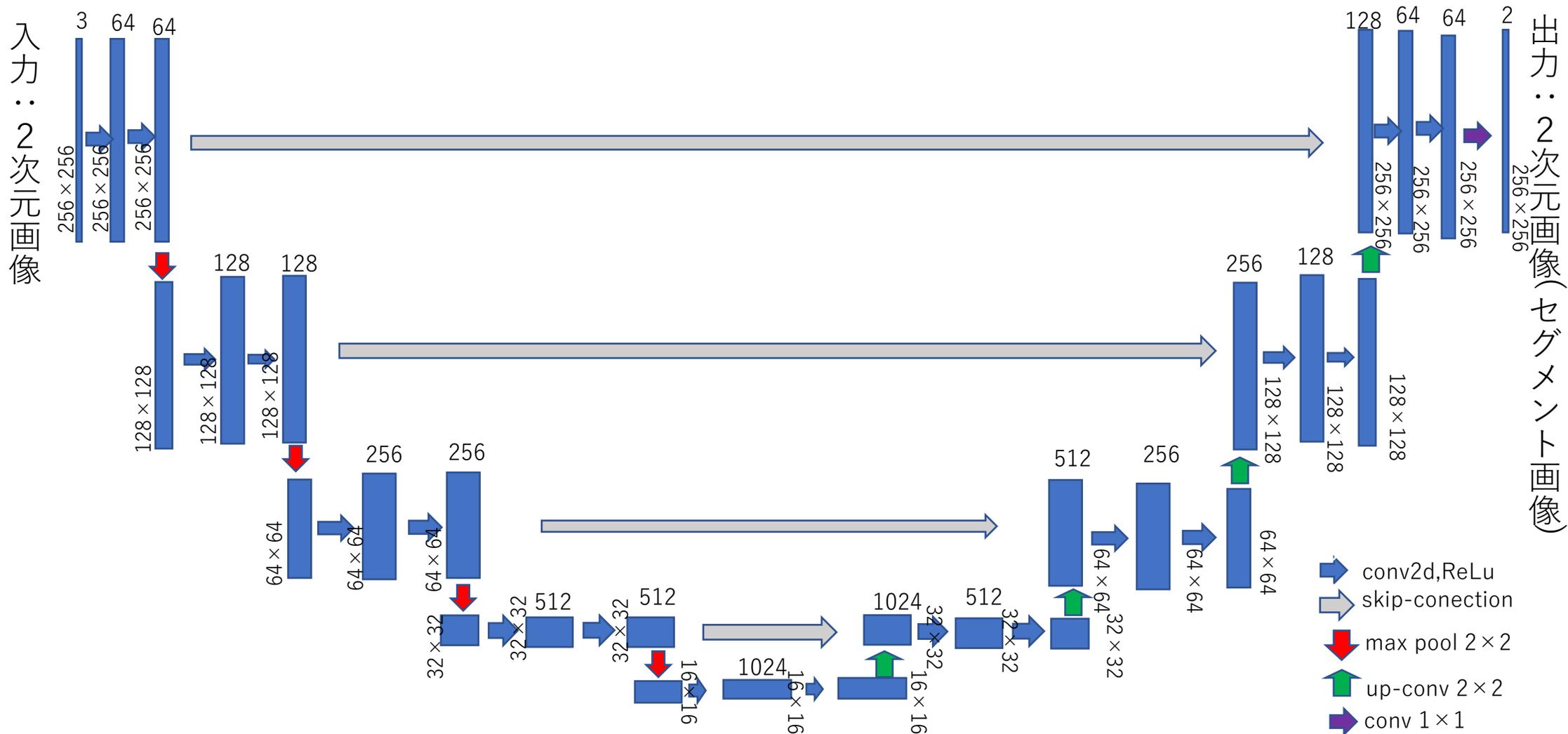


入力画像



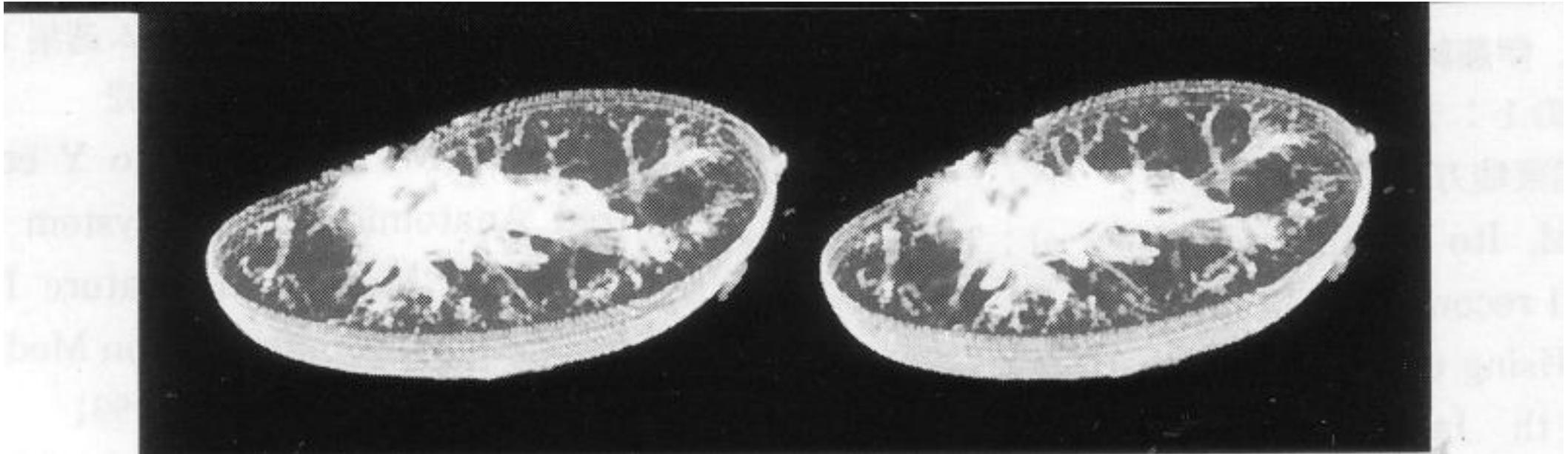
予測画像
(セグメント画像)

U-Netのネットワーク構造



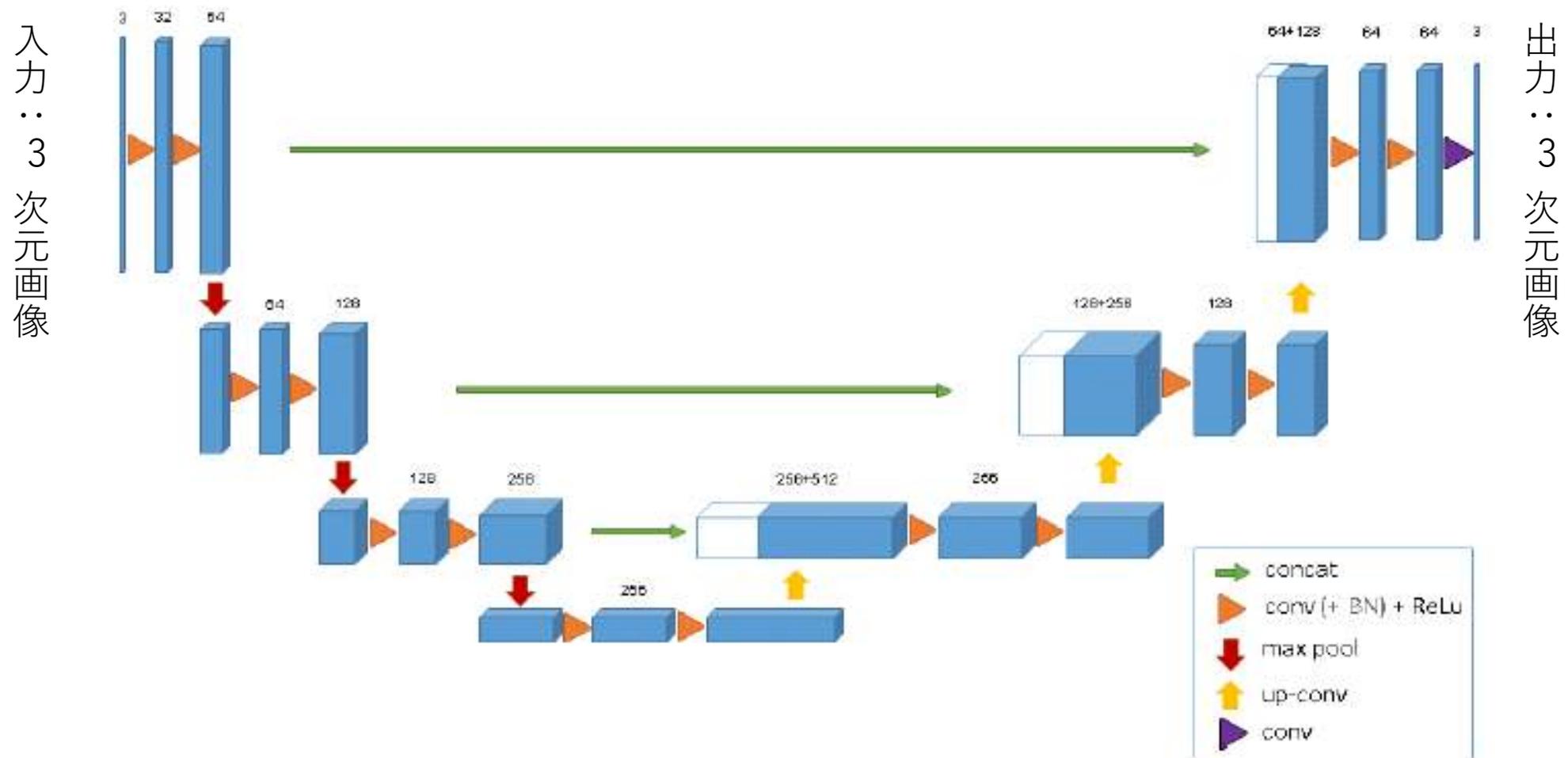
従来研究②

- ボリュームデータ [8] への適用 → U-Net3D [2]



[8] 鈴木雅隆, 柴田昌和, 周藤安造, “ボリューム・レンダリングの解剖学への応用 ラット腎臓の連続組織切片の3次元構築”, MEDICAL IMAGING TECHNOLOGY Vol.13 No.3 May(1995)

U-Net3Dのネットワーク構造[2]



提案手法

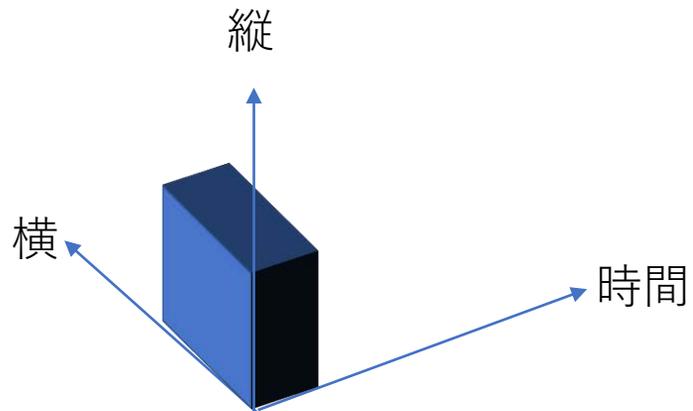
時間方向に束ねて、3次元画像としてU-Net3Dに適用



時間方向と空間方向では性質が異なる

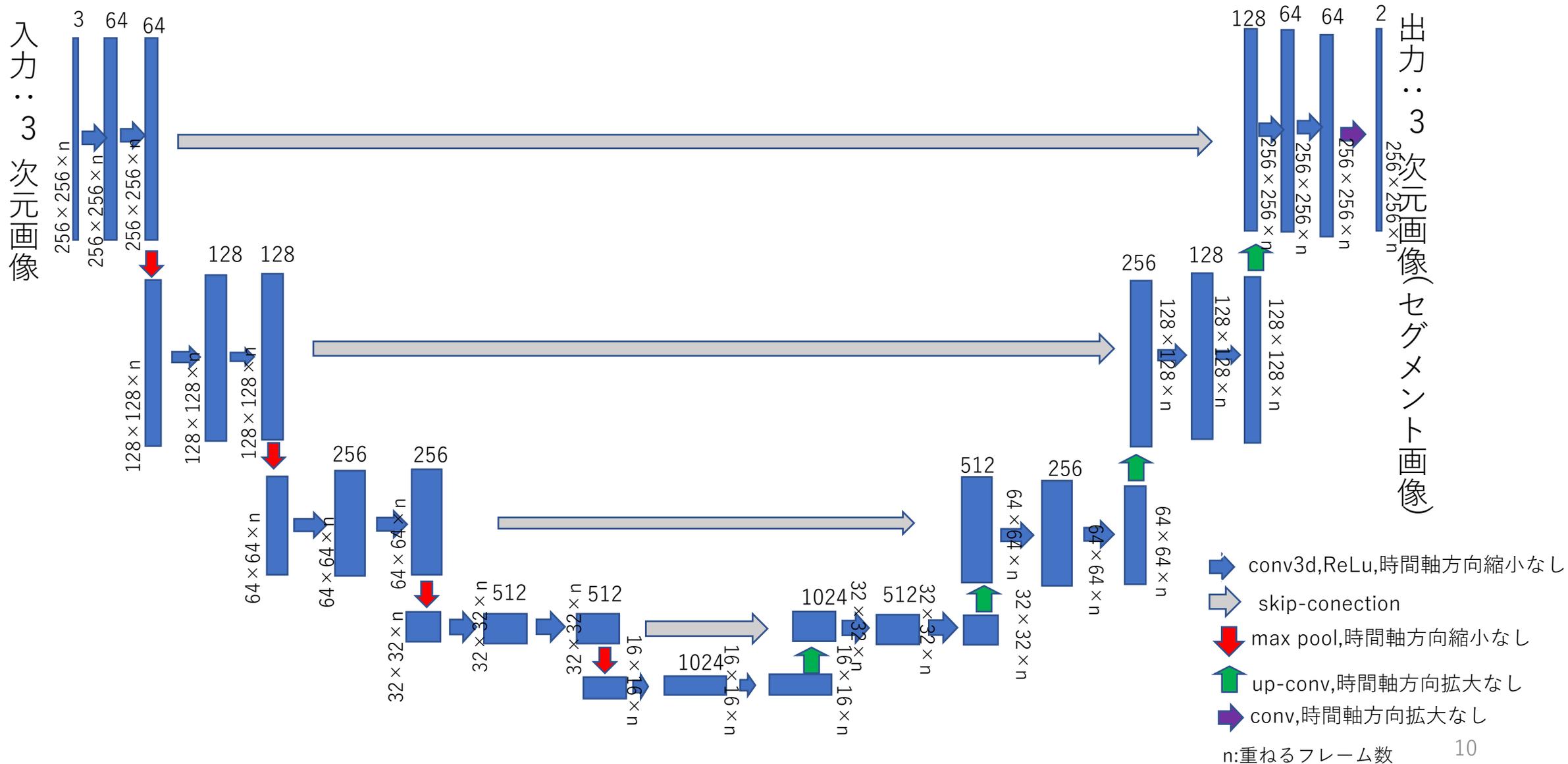


U-Net3DTを提案



※ 3次元画像：クリップ

U-Net3DTのネットワーク構造



実験

- U-Net,U-Net3D,U-Net3DTの比較

	学習データ	テストデータ
test1	• CIMG0120(黒牛 1頭) フレーム数 256枚	• CIMG0120(黒牛 1頭) フレーム数 98枚
test2	• test1と同じ	• CIMG0123(黒牛 1頭 白牛 1頭) フレーム数 216枚
test3	• test1と同じ	• MVI_357(黒牛 2頭) フレーム数 280枚



CIMG0120



CIMG0123

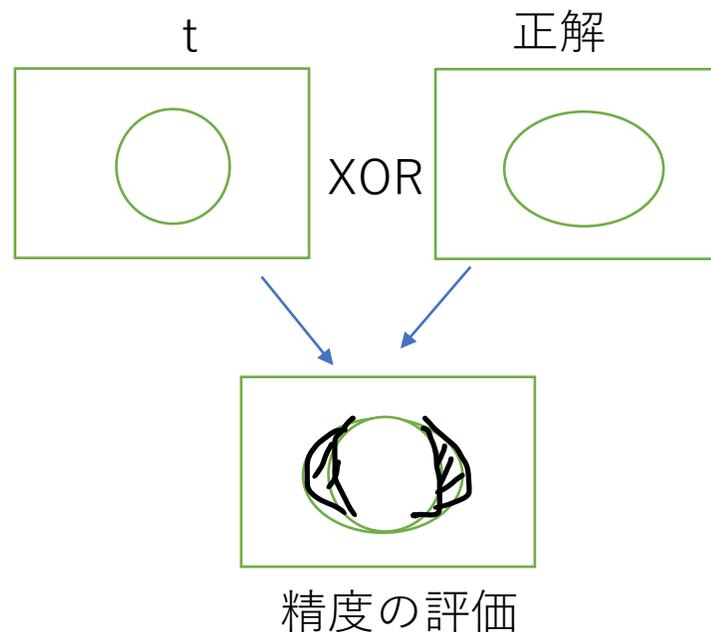
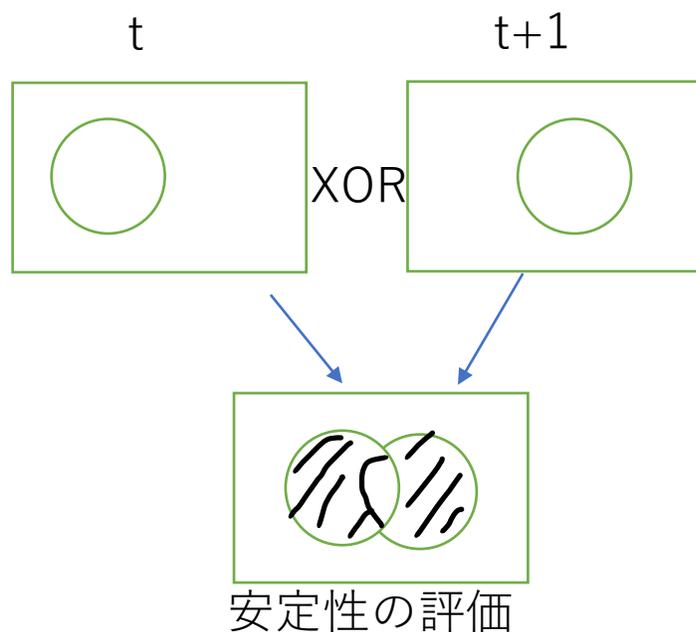


MVI_357

評価方法

異なる画素値の総数を集計

- **安定性の評価**：連続した隣同士の画像を比較
- **精度の評価**：正解画像と比較



test1,2の出力結果



元画像



2D

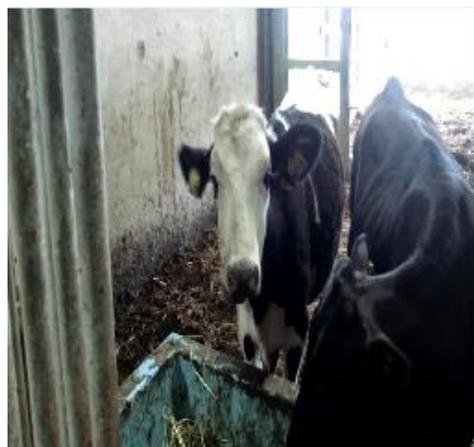


3D



3DT

test1



元画像



2D



3D



3DT

test2

test3の出力結果



元画像



2D



3D

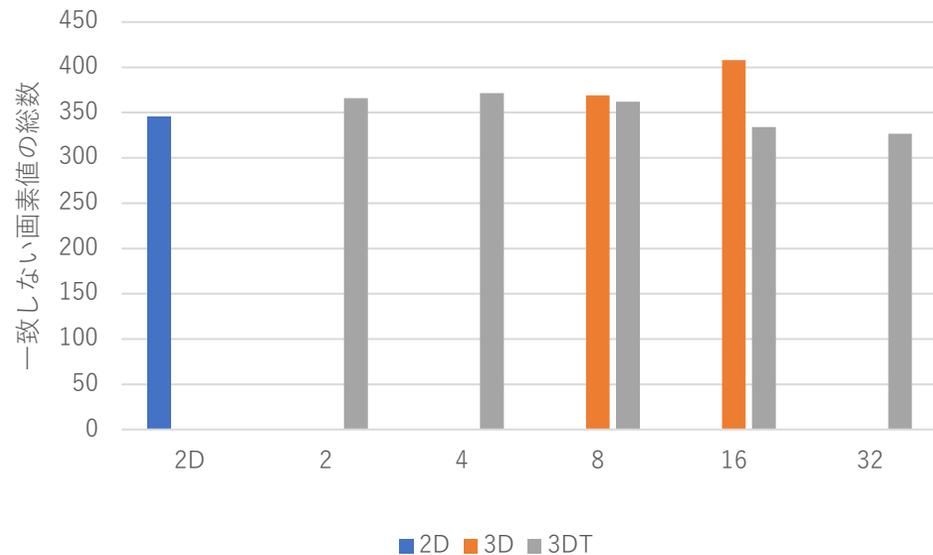


3DT

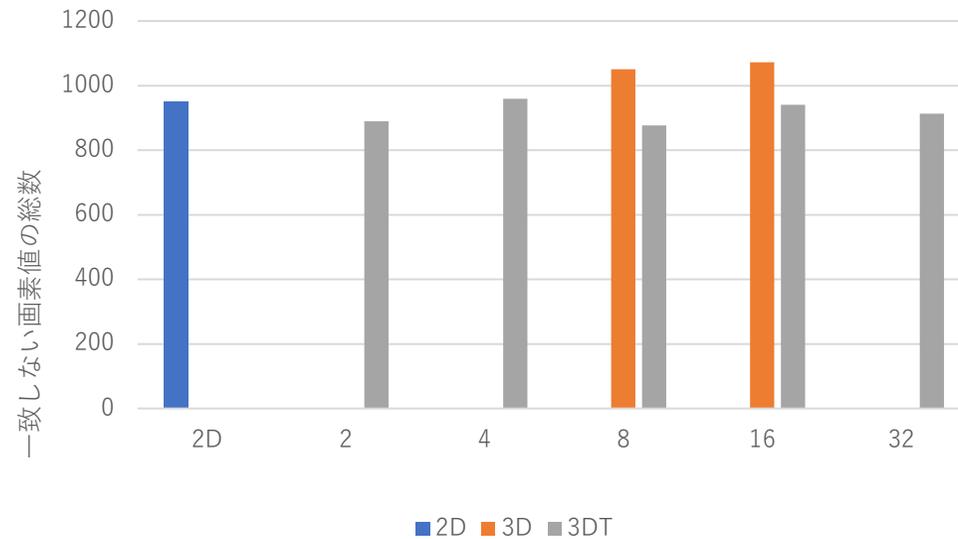
test3

結果と考察①

- ほぼ誤差の範囲
- 2D,3D,3DTのすべて検出が上手くいっている
→学習データとテストデータが似ていたから



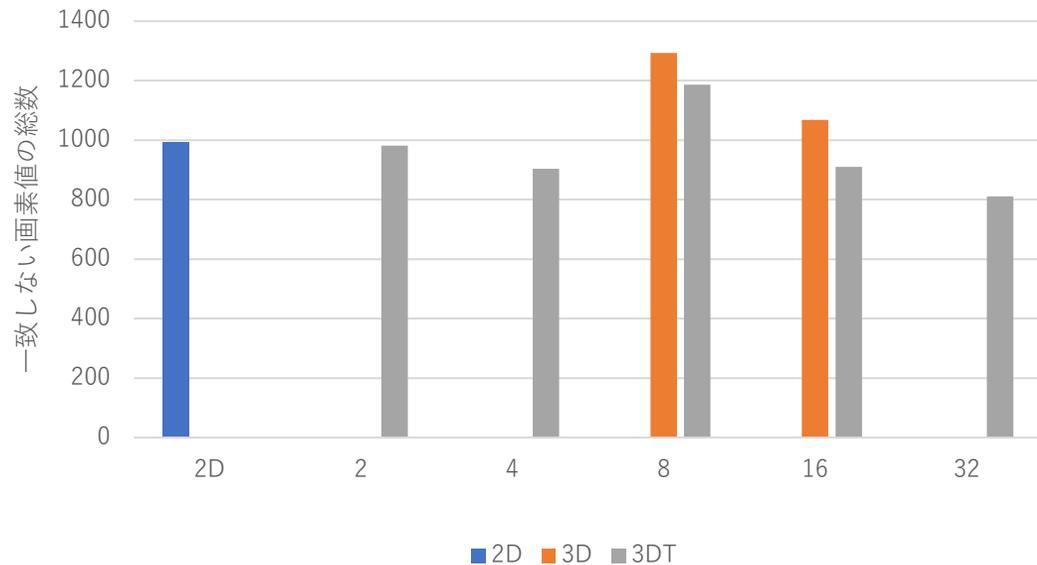
(a)test1 隣比較



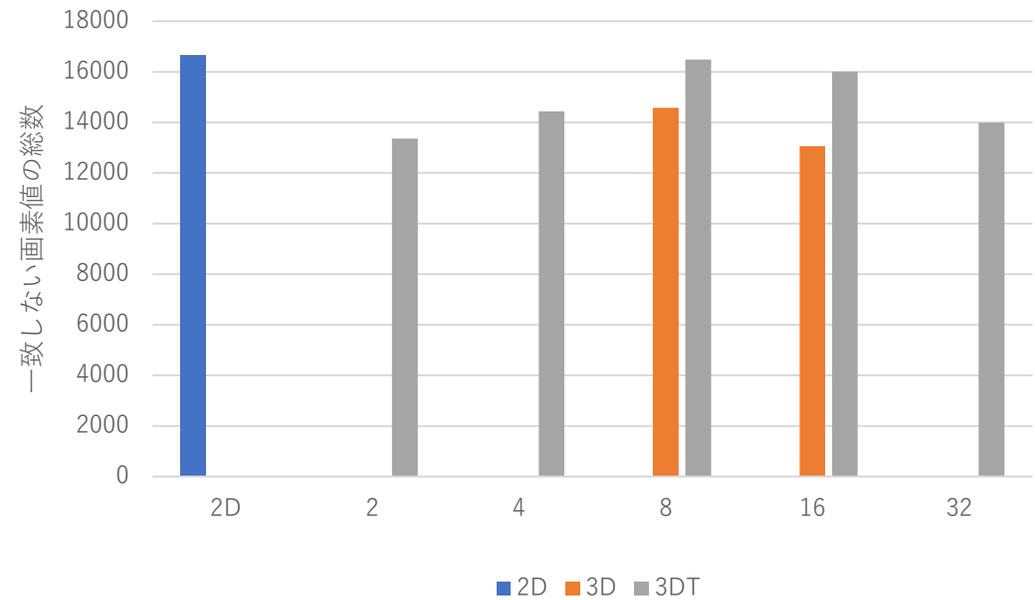
(b)test1 正解比較

結果と考察②

- 全体的な傾向はみられない。
→学習データとテストデータ大きな違いがあったから



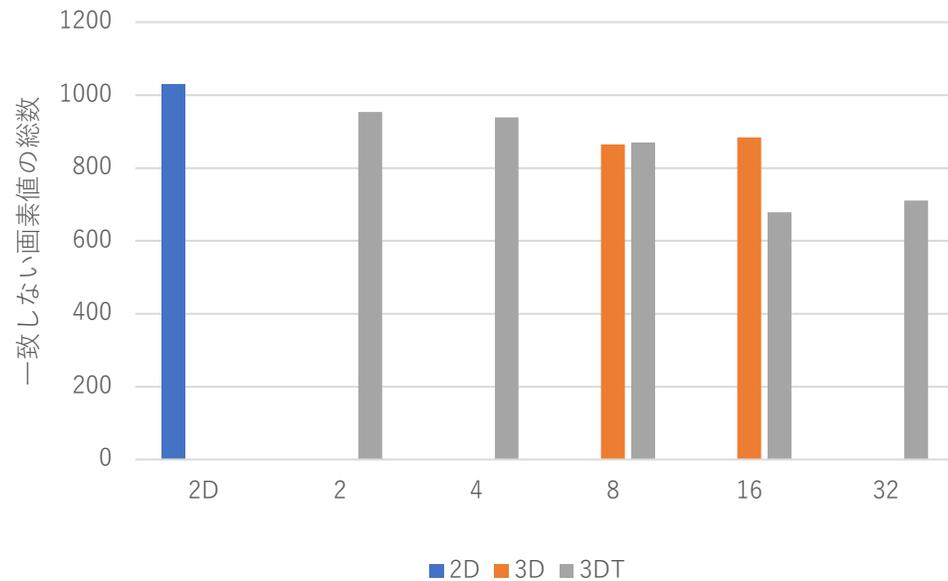
(a) test2 隣比較



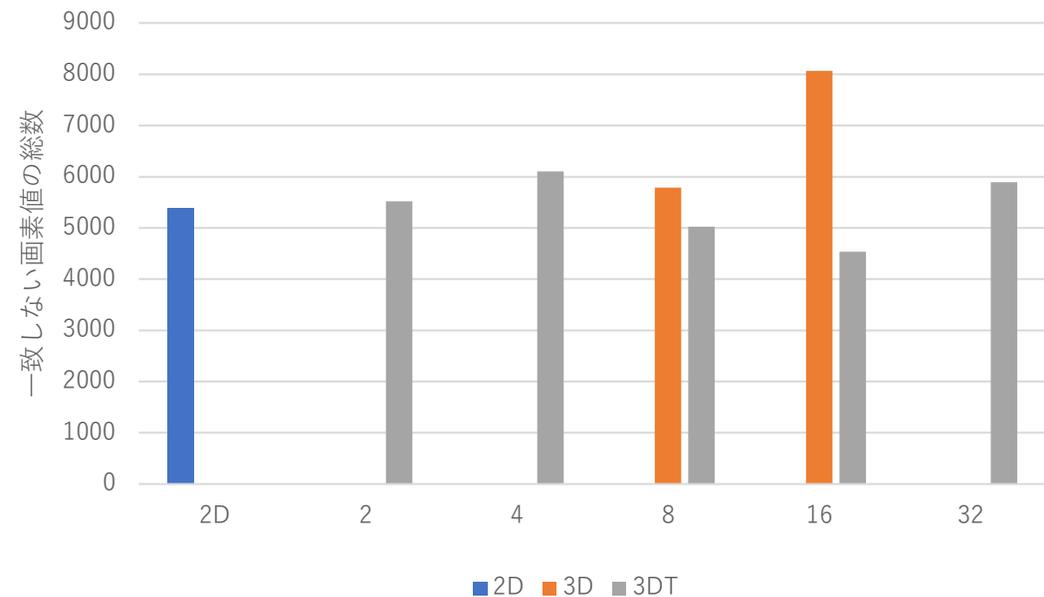
(b) test2 正解比較

結果と考察③

- クリップの長さによっては3DTの良い値が出たと考えられる。
→学習データとテストデータが似ていたから



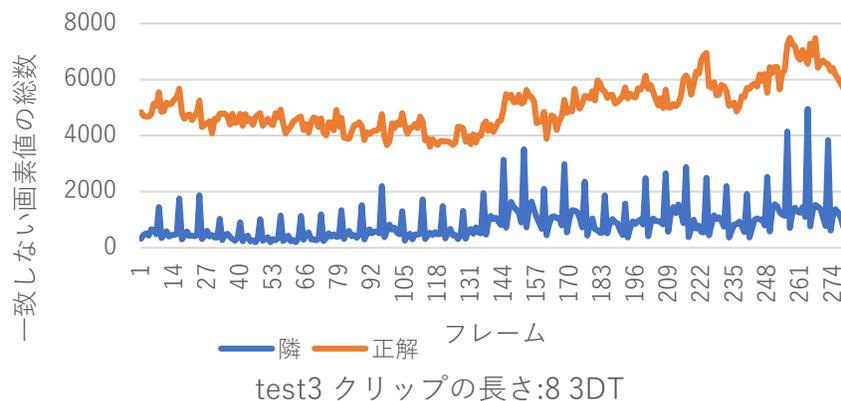
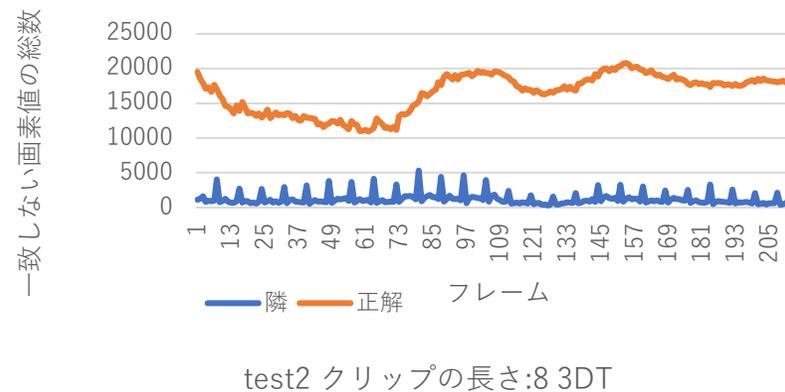
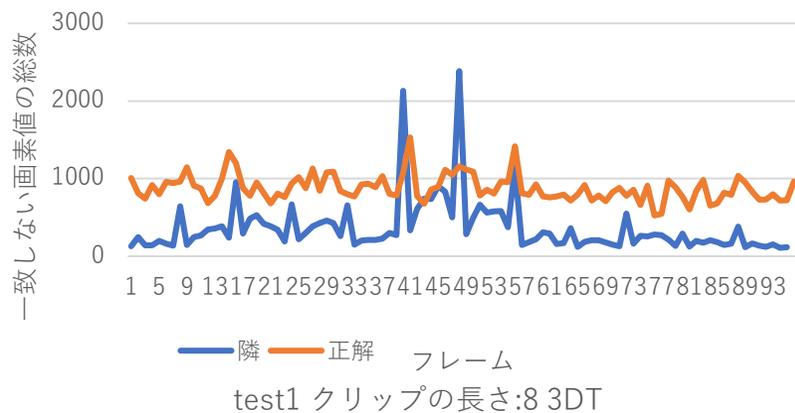
(a) test3 隣比較



(b) test3 正解比較

結果と考察④

- 8の倍数（クリップの長さ）で物体検出の精度が落ちる。
→クリップの長さ分だけ異なる動きが出ているから



まとめ

- 連続検出の安定化のためにU-Net3DTを提案
- 提案手法（U-Net3DT）と従来手法（U-Net,U-Net3D）を比較



提案手法は安定性、精度の点で優れていた。

今後の課題

- 追加実験
- さらなる安定化、高精度化に向けた検討

参考文献

[2] Özgün Çiçek, Ahmed Abdulkadir, Soeren S. Lienkamp, Thomas Brox, Olaf Ronneberger, “3D U-Net: Learning Dense Volumetric Segmentation from Sparse Annotation”, arXiv:1606.06650 [cs.CV](2016)