

個体毎の領域分割を用いた 遮蔽に強い複数物体追跡

工学研究科 情報システム工学分野

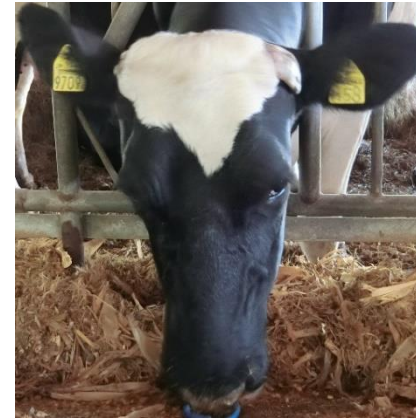
T1903290 竹井 奏一郎

指導教員 椋木 雅之

2021/02/01

研究背景

- 牛の中には要注意な個体が存在する
伝染病を持っている可能性のある個体



- 牛の個体毎に行動追跡する技術は重要
 - タグ：汚れや顔の向きで判断できない場合がある
 - 発信機[1]：機器の装着による牛への負担
 - 広範囲で牛が移動する場合、受信設備等の負担

広範囲を撮影した映像から牛を追跡できれば個体管理に役立つ
→ **物体追跡技術**を利用

[1]ファームノート <https://farmnote.jp/press-release/farmnote-color2016.html>

研究目的

動画中の複数物体（牛）追跡

- インスタンス・セグメンテーションを利用した個体検出
- 検出結果を時間方向に対応付けた追跡

牛の個体管理のために高精度な追跡の実現

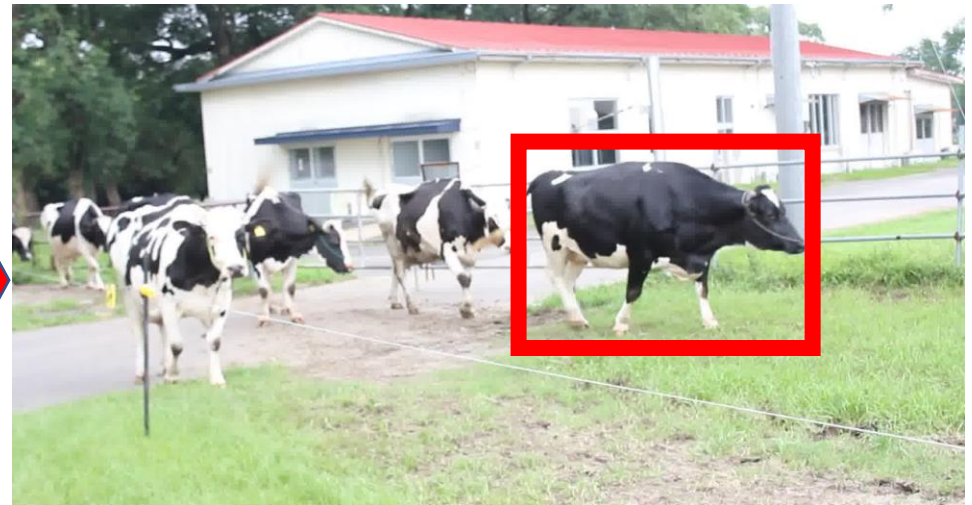
物体追跡問題

動画像中の指定した位置に存在する物体が、
どのように移動したか推定する問題

(本研究では特に飼育環境下の牛を対象とした物体追跡)



時間： t_n



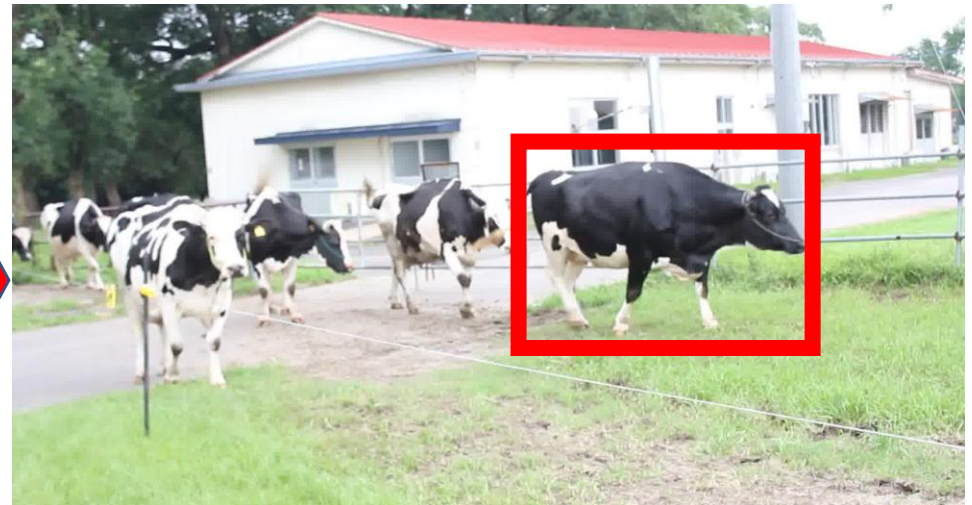
時間： t_{n+1}

従来の物体追跡手法

- 先頭フレームで指定した情報を用いる
 - バウンディングボックスで指定
- 指定した情報と類似した画像領域を時間方向に対応付けて追跡
(追跡対象の見え方の事前知識なし)



時間： t_n



時間： t_{n+1}

従来の物体追跡の問題点

遮蔽が生じると追跡に失敗する

遮蔽：手前の物体が後方の物体を隠す

- ・ 追跡対象の一部が見えなくなる（見え方が変わる）
- ・ 類似個体同士が重なる（個体毎の検出が困難）



時間： t_1



...



時間： t_n

従来の物体追跡②

Tracking by Detection：検出に基づく追跡

- 追跡対象の見え方の事前知識あり
- 検出結果を時間方向に対応付けて追跡

→個体毎の見え方の変化にも対応
(隣接や重なりなど)

検出の精度が高くなければ追跡ができない

ディープラーニングを用いた物体検出技術

検出の精度が高い

- 一部しか見えない物体の検出が可能
- 複数個体が隣接していても個体毎の検出が可能

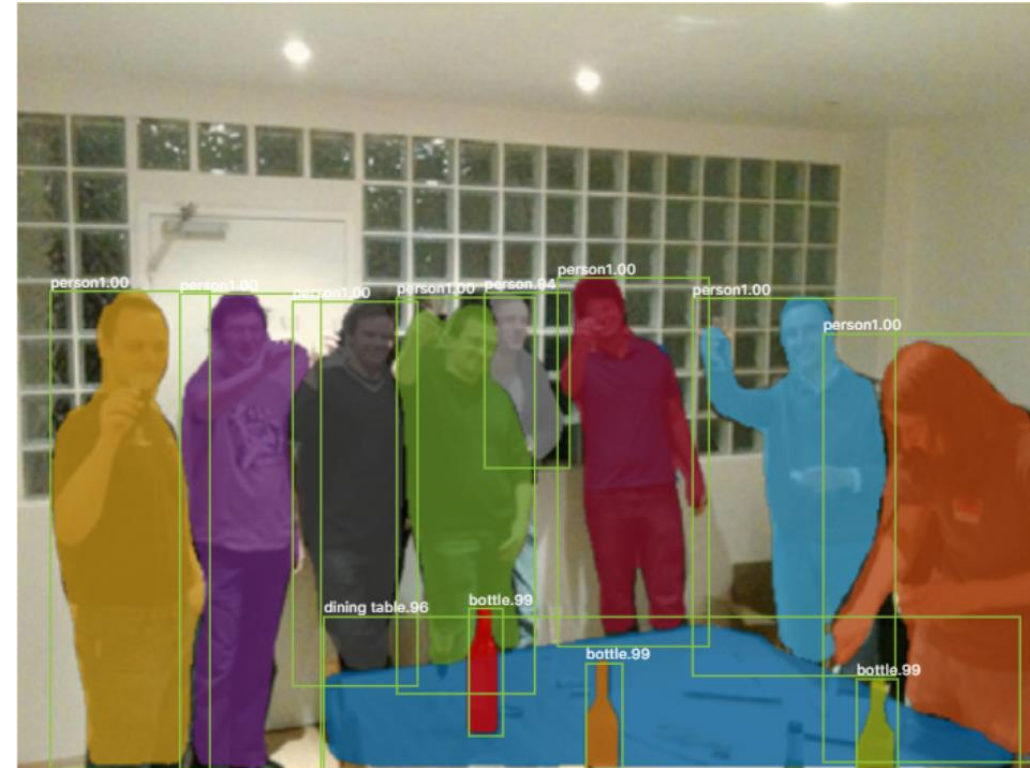
本研究：

インスタンス・セグメンテーションを利用

インスタンス・セグメンテーション

入力された画像内の個体毎の領域をピクセルレベルで検出

- 遮蔽があっても検出可能
- 正確な領域を抽出



提案手法

インスタンス・セグメンテーションを用いた個体毎の検出
(精度が高く、遮蔽に強い)

+

検出結果を時間方向に対応付けた追跡

||

精度の高い物体検出結果に基づいた複数物体追跡

提案手法の処理手順

処理手順

- ① 追跡対象の個体数と先頭フレームでの位置の指定
- ② インスタンス・セグメンテーションによる候補領域の検出
- ③ 候補領域の組み合わせの作成
- ④ 現在フレームと次フレームにおける、追跡対象と組み合わせた候補の重なるの算出
- ⑤ 追跡対象と候補領域の対応付け
- ⑥ ③～⑤の繰り返し

処理手順

- ① **追跡対象の個体数と先頭フレームでの位置の指定**
- ② インスタンス・セグメンテーションによる候補領域の検出
- ③ 候補領域の組み合わせの作成
- ④ 現在フレームと次フレームにおける、追跡対象と組み合わせた候補の重なるの算出
- ⑤ 追跡対象と候補領域の対応付け
- ⑥ ③～⑤の繰り返し

追跡対象の個体数と 先頭フレームでの位置の指定

個体数：2

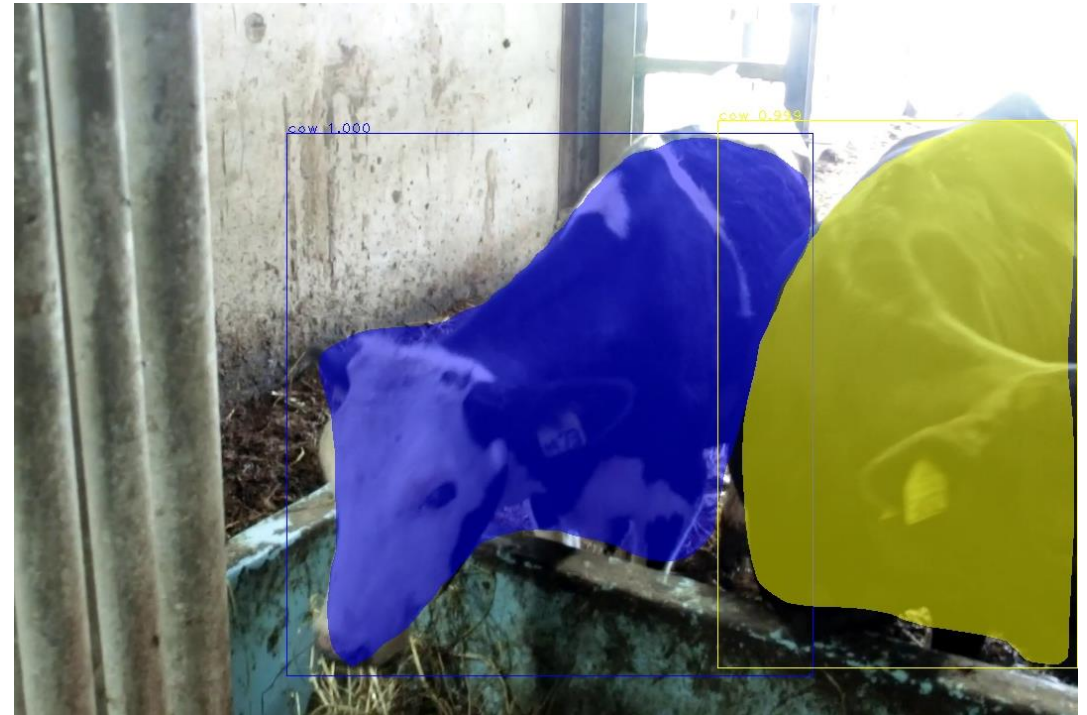


処理手順

- ① 追跡対象の個体数と先頭フレームでの位置の指定
- ② **インスタンス・セグメンテーションによる候補領域の検出**
- ③ 候補領域の組み合わせの作成
- ④ 現在フレームと次フレームにおける、追跡対象と組み合わせた候補の重なるの算出
- ⑤ 追跡対象と候補領域の対応付け
- ⑥ ③～⑤の繰り返し

候補領域の検出

- インスタンス・セグメンテーションを実行するモデル：
Mask R-CNN
- 検出結果：マスク領域



候補領域の検出時の問題①

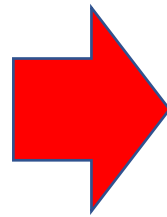
1つの個体が過剰に分割され検出される場合がある



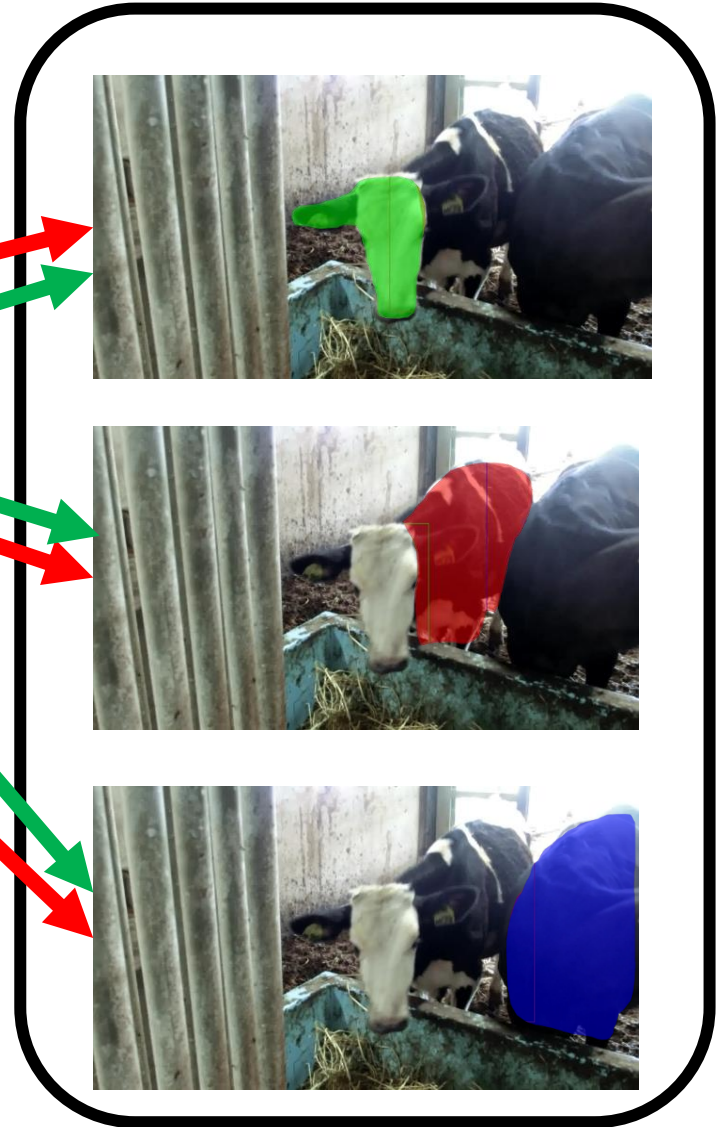
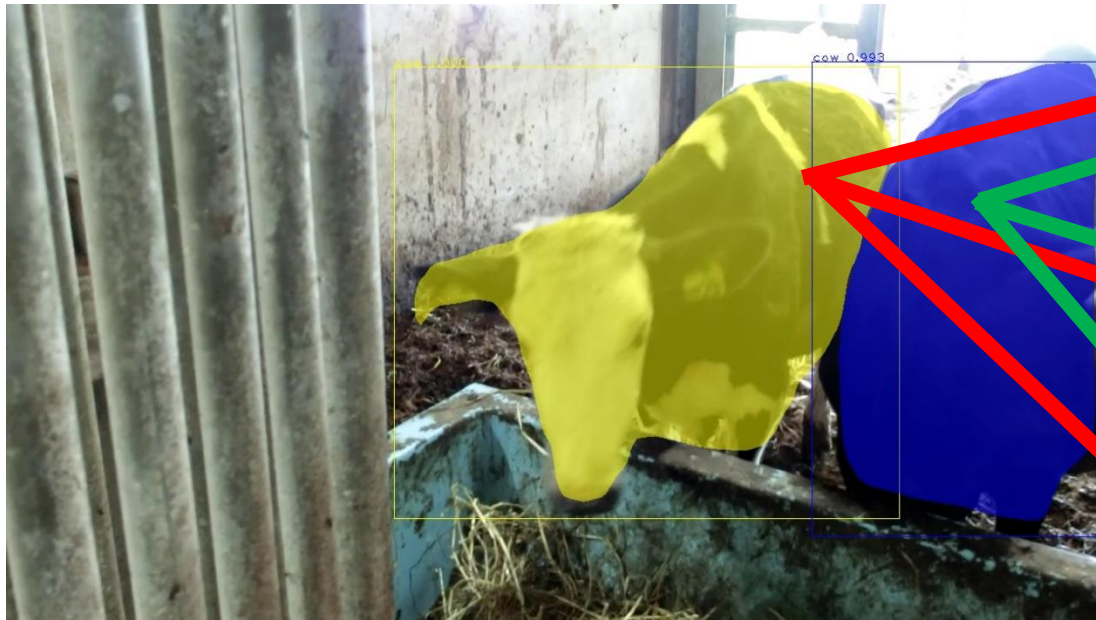
処理手順

- ① 追跡対象の個体数と先頭フレームでの位置の指定
- ② インスタンス・セグメンテーションによる候補領域の検出
- ③ 候補領域の組み合わせの作成**
- ④ 現在フレームと次フレームにおける、追跡対象と組み合わせた候補の重なるの算出
- ⑤ 追跡対象と候補領域の対応付け
- ⑥ ③～⑤の繰り返し

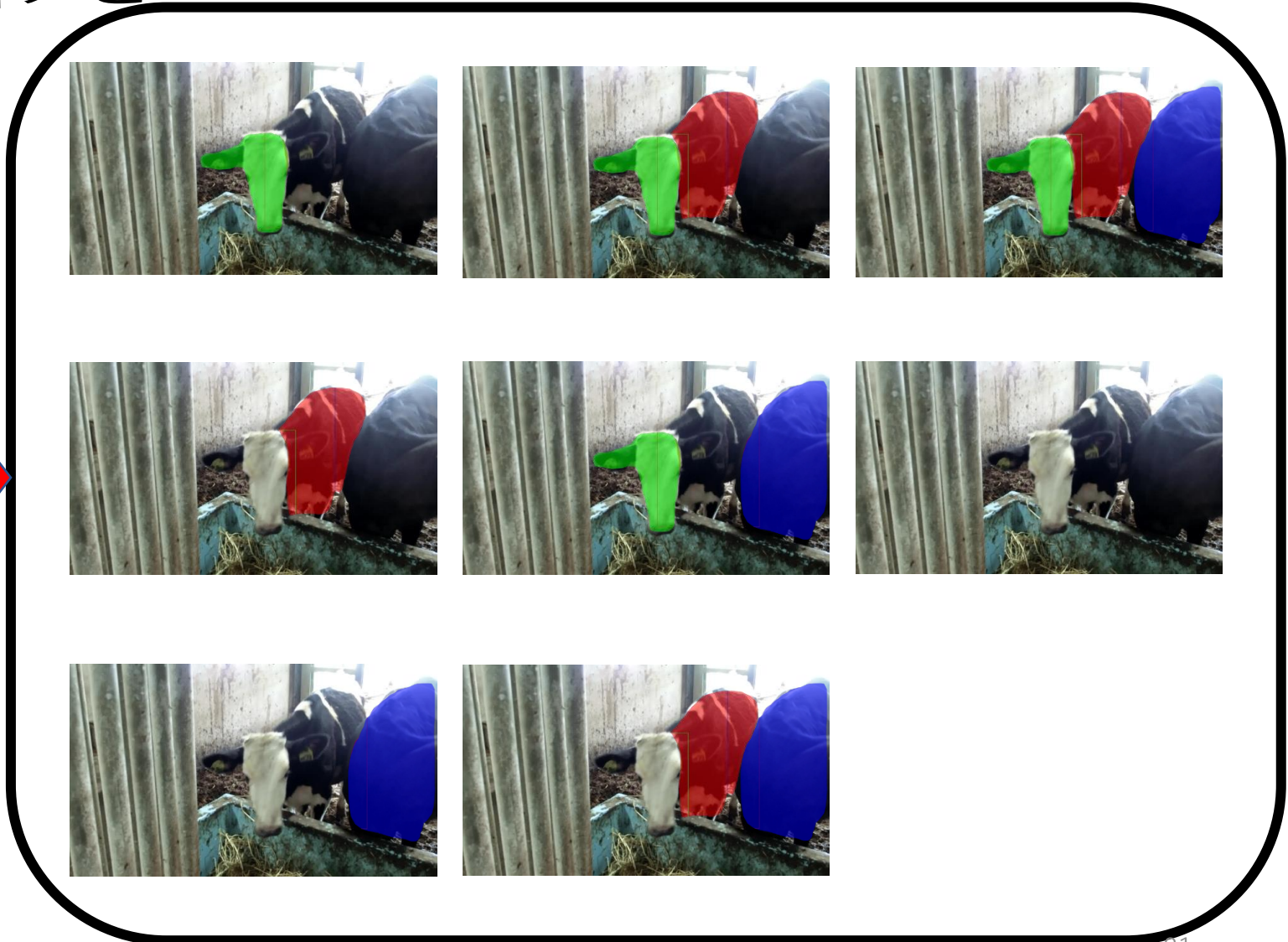
候補の組み合わせ



候補の組み合わせ



候補の組み合わせ



候補の組み合わせ

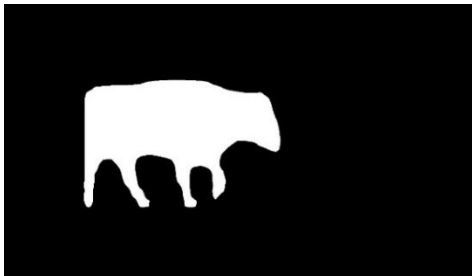


処理手順

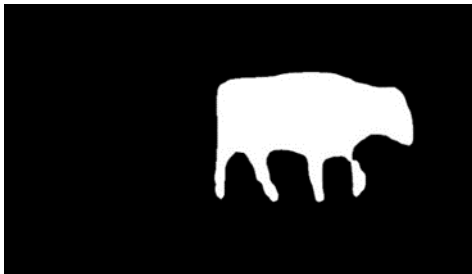
- ① 追跡対象の個体数と先頭フレームでの位置の指定
- ② インスタンス・セグメンテーションによる候補領域の検出
- ③ 候補領域の組み合わせの作成
- ④ **現在フレームと次フレームにおける、追跡対象と組み合わせた候補の重なるの算出**
- ⑤ 追跡対象と候補領域の対応付け
- ⑥ ③～⑤の繰り返し

重なるの計算

- IoU (Intersection over Union)



現在フレームの
追跡対象の領域



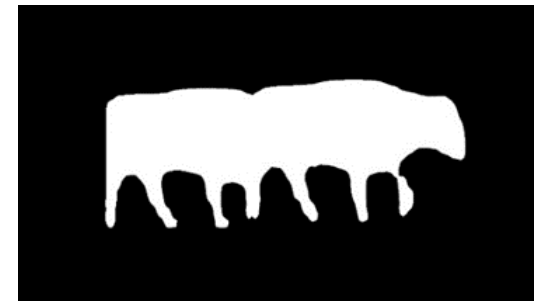
次フレームの
候補領域

重複領域



$$\text{IoU} = \frac{\text{重複領域}}{\text{全体領域}}$$

全体領域



処理手順

- ① 追跡対象の個体数と先頭フレームでの位置の指定
- ② インスタンス・セグメンテーションによる候補領域の検出
- ③ 候補領域の組み合わせの作成
- ④ 現在フレームと次フレームにおける、追跡対象と組み合わせた候補の重なるの算出
- ⑤ 追跡対象と候補領域の対応付け**
- ⑥ ③～⑤の繰り返し

追跡対象と候補領域の対応付け

追跡対象 α : 次フレーム候補 $\{a, b, a+b, \varepsilon\}$

追跡対象 β : 次フレーム候補 $\{a, b, a+b, \varepsilon\}$

ε : 空要素



追跡対象と候補領域の対応付け

追跡対象 α : 次フレーム候補 $\{\textcircled{a}, b, a+b, \varepsilon\}$

追跡対象 β : 次フレーム候補 $\{a, \textcircled{b}, a+b, \varepsilon\}$

ε : 空要素



追跡対象と候補領域の対応付け

追跡対象 α : 次フレーム候補 $\{a, b, a+b, \varepsilon\}$

追跡対象 β : 次フレーム候補 $\{a, b, a+b, \varepsilon\}$

ε : 空要素

同じ要素の候補の使用は許容しない



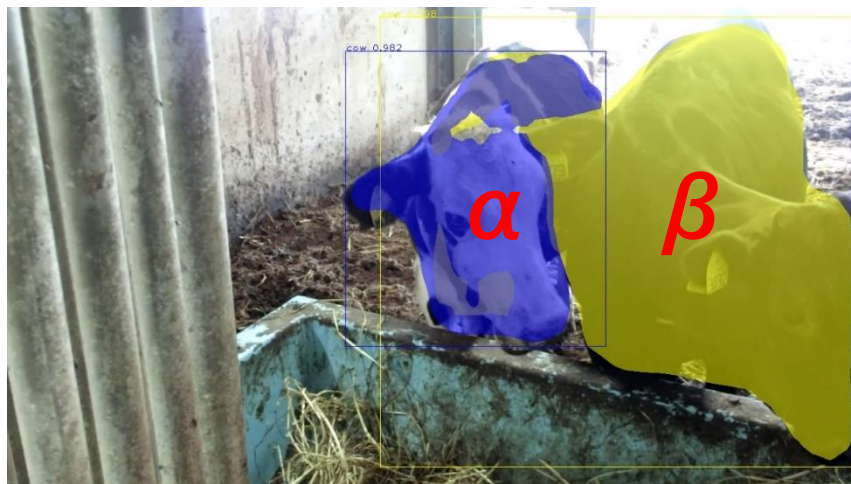
追跡対象と候補領域の対応付け

追跡対象 α : 次フレーム候補 $\{a, b, a+b, \varepsilon\}$

追跡対象 β : 次フレーム候補 $\{a, b, a+b, \varepsilon\}$

ε : 空要素

IoU値の和が最大の組み合わせを採用



候補領域の検出時の問題②

一部の個体が検出されない検出漏れが生じる



追跡対象と候補領域の対応付け

追跡対象 α : 次フレーム候補 $\{b, \varepsilon\}$

追跡対象 β : 次フレーム候補 $\{b, \varepsilon\}$

ε : 空要素

追跡対象は移動していないものとして処理



処理手順

• 提案手法の処理手順

- ① 追跡対象の個体数と先頭フレームでの位置の指定
- ② インスタンス・セグメンテーションによる候補領域の検出
- ③ 候補領域の組み合わせの作成
- ④ 現在フレームと次フレームにおける、追跡対象と組み合わせた候補の重なるの算出
- ⑤ 追跡対象と候補領域の対応付け
- ⑥ **③～⑤の繰り返し**

評価実験

実験設定

- 比較手法

従来手法：OpenCVに実装されている7つの手法

- インスタンス・セグメンテーションに用いるMask R-CNN
牛を含む20種類のクラスが識別できるように学習済み

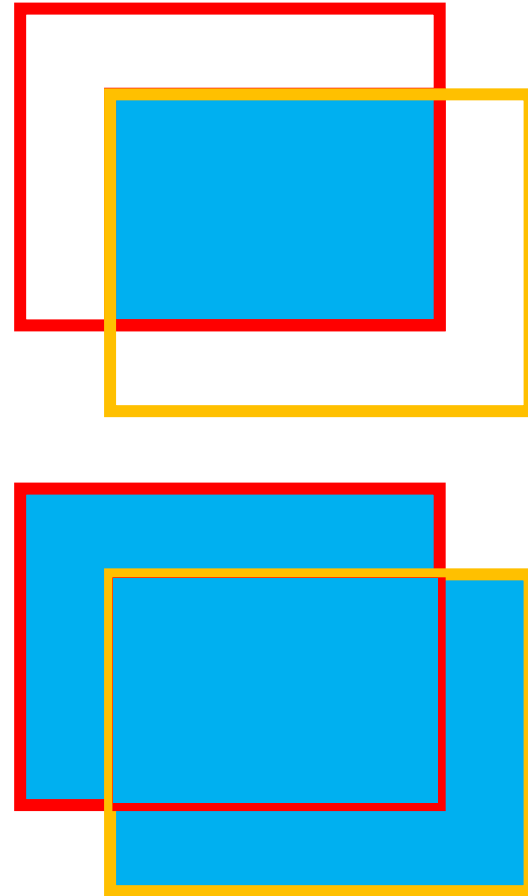
- 評価指標

IoU、重心位置の平均誤差

評價指標

IoU (Intersection over Union)

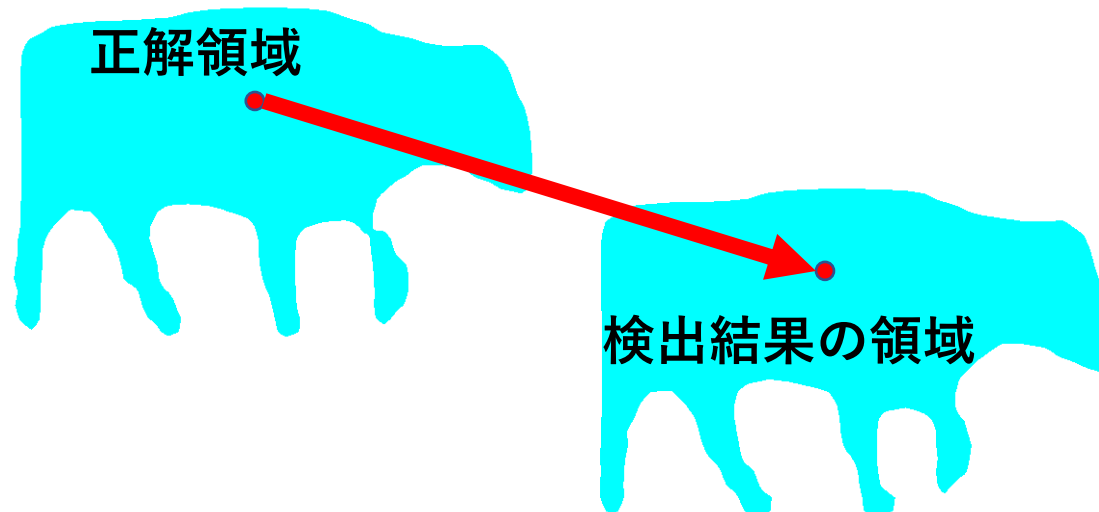
$$\bullet \text{IoU} = \frac{\text{重複領域}}{\text{全体領域}}$$



評価指標

重心位置の平均誤差

正解領域の重心と検出結果の重心のユークリッド距離



実験データ

動画名	個体数	フレーム数	撮影環境	移動状態	牛の色	特徴
C123	2	216	牛舎内	あまり移動しない	白と黒	後方に光源
M358	2	343	牛舎内	あまり移動しない	茶	柵による遮蔽 常時牛同士の重なり
M724	7	182	屋外	大きく移動する	白と黒	遠方からの撮影 奥から手前へ進行

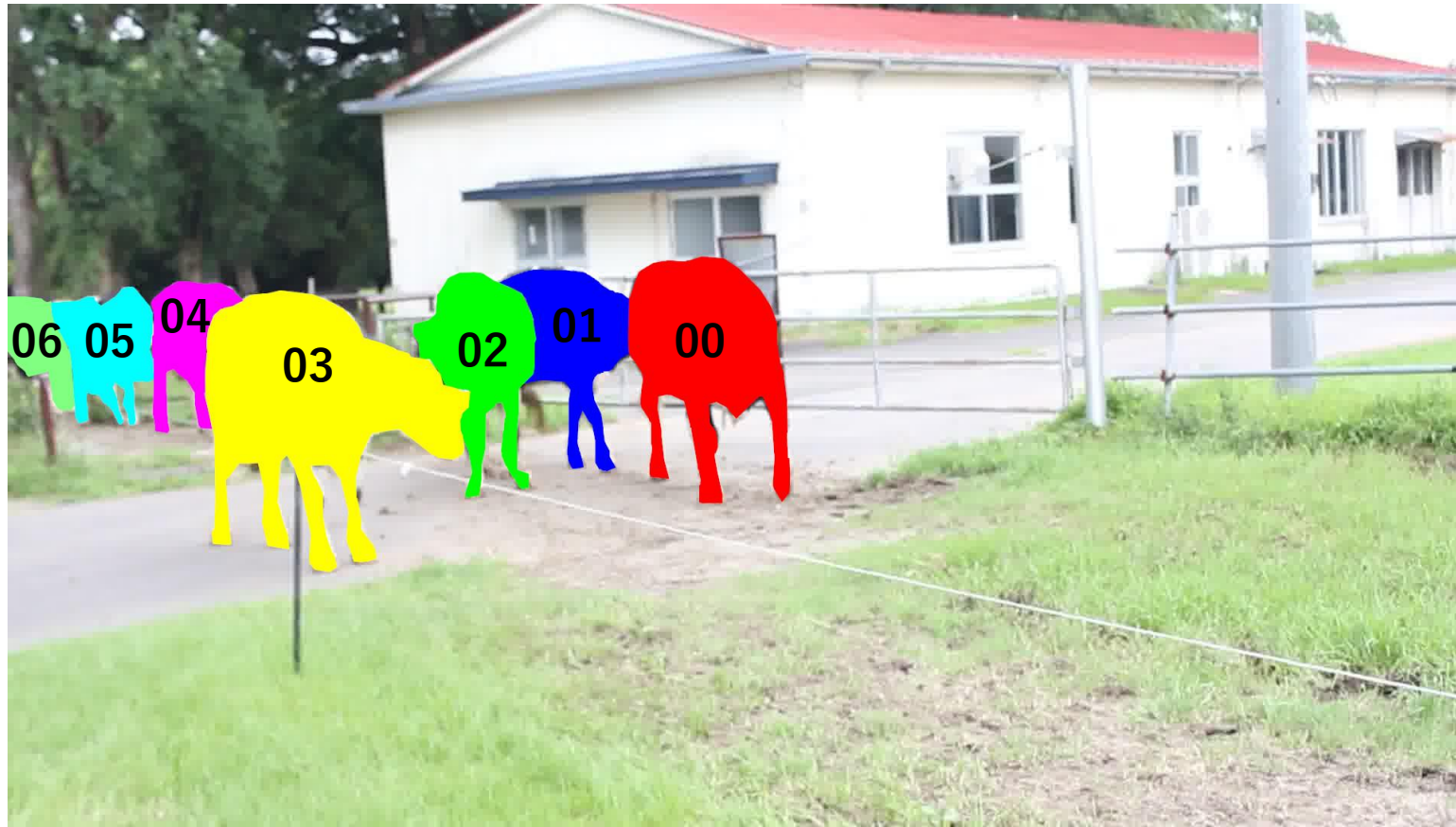
M724：画像例

- 右から順番にかけて追跡対象を区別



M724：正解データ

- 右から順番にかけて追跡対象を区別

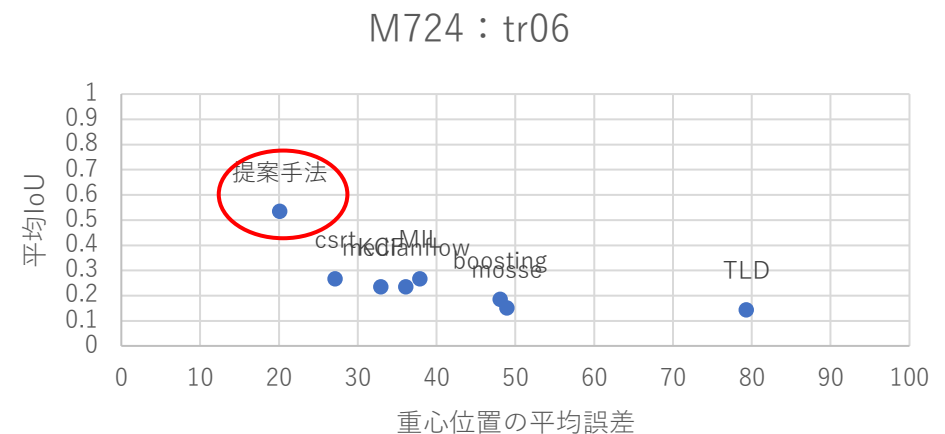
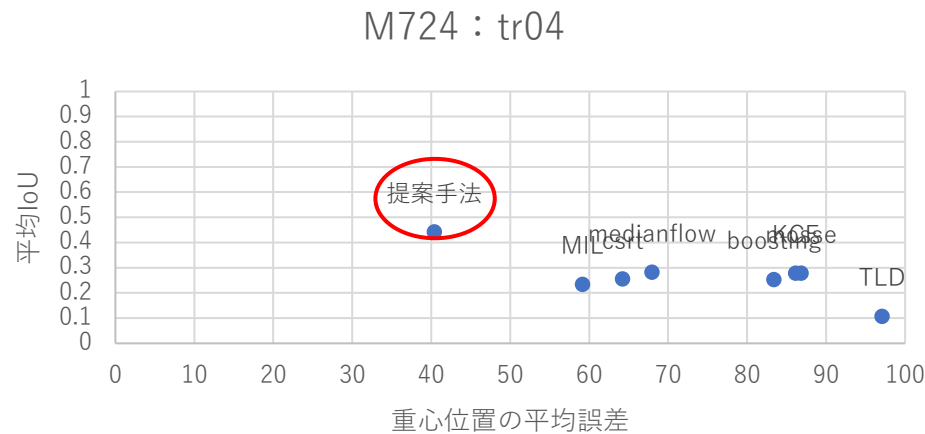
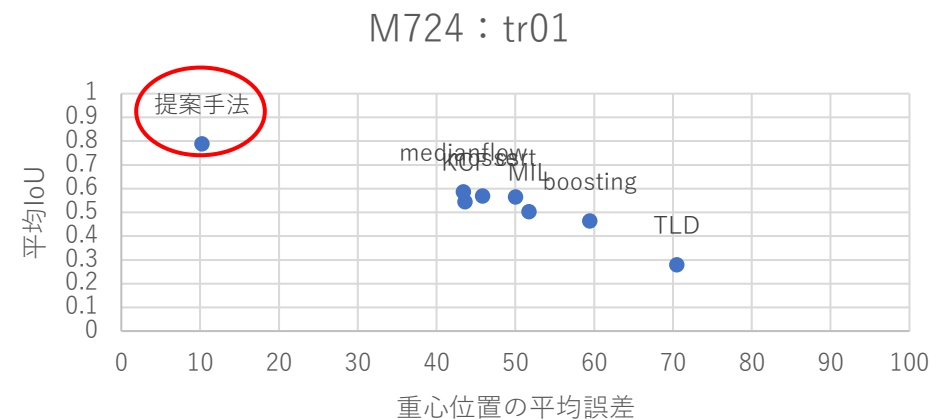
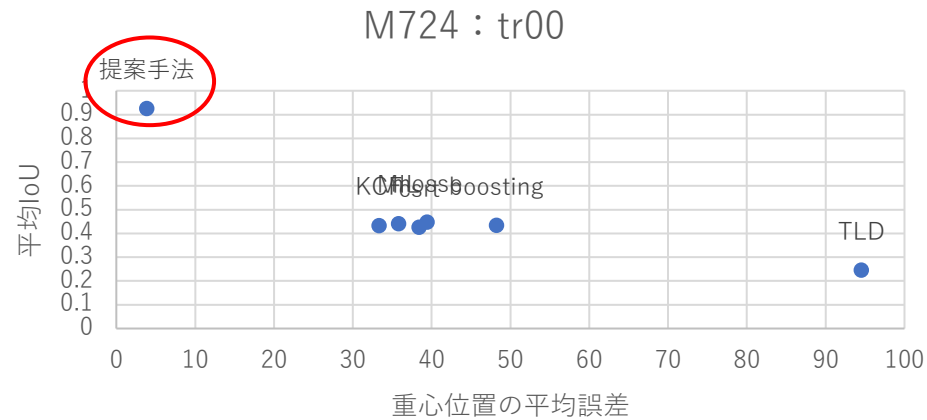


実験結果 (M724)



実験結果 (M724)

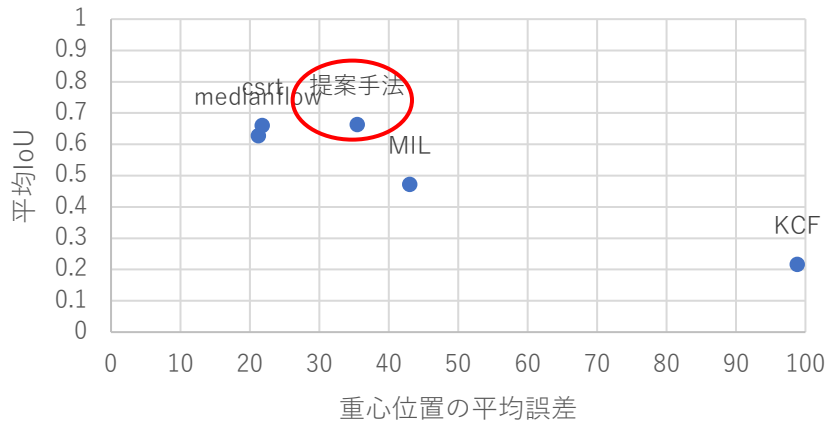
良好な結果：tr00, tr01, tr04, tr06



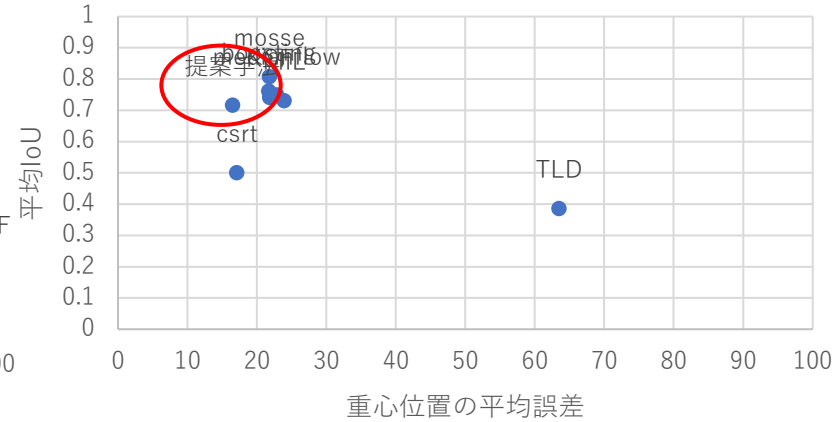
実験結果 (M724)

従来手法と同等：tr02, tr03, tr05

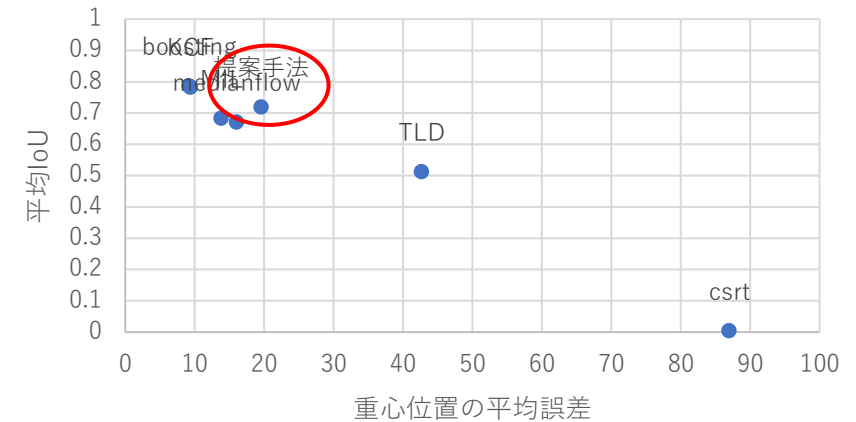
M724 : tr02



M724 : tr03

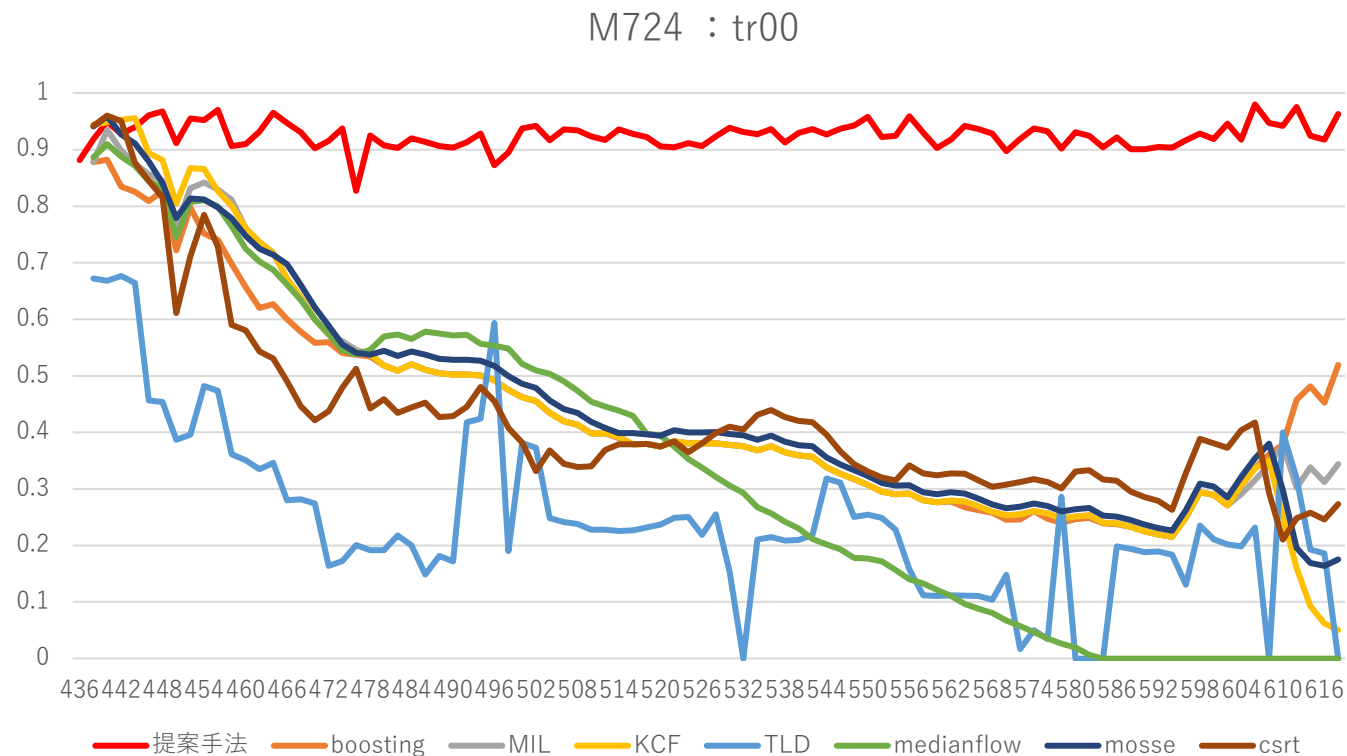


M724 : tr05



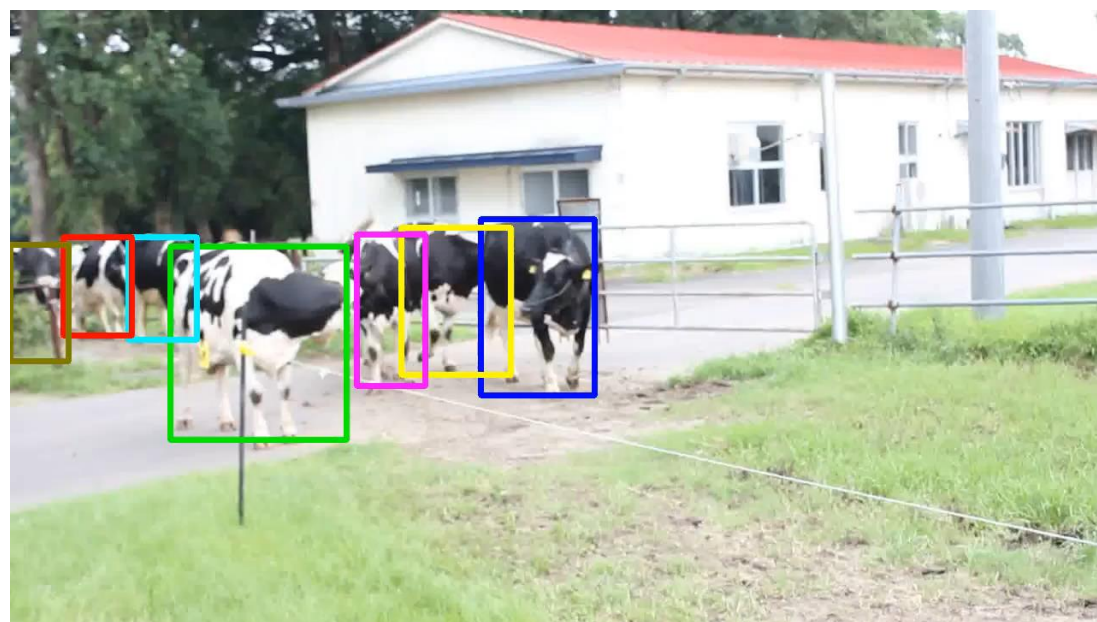
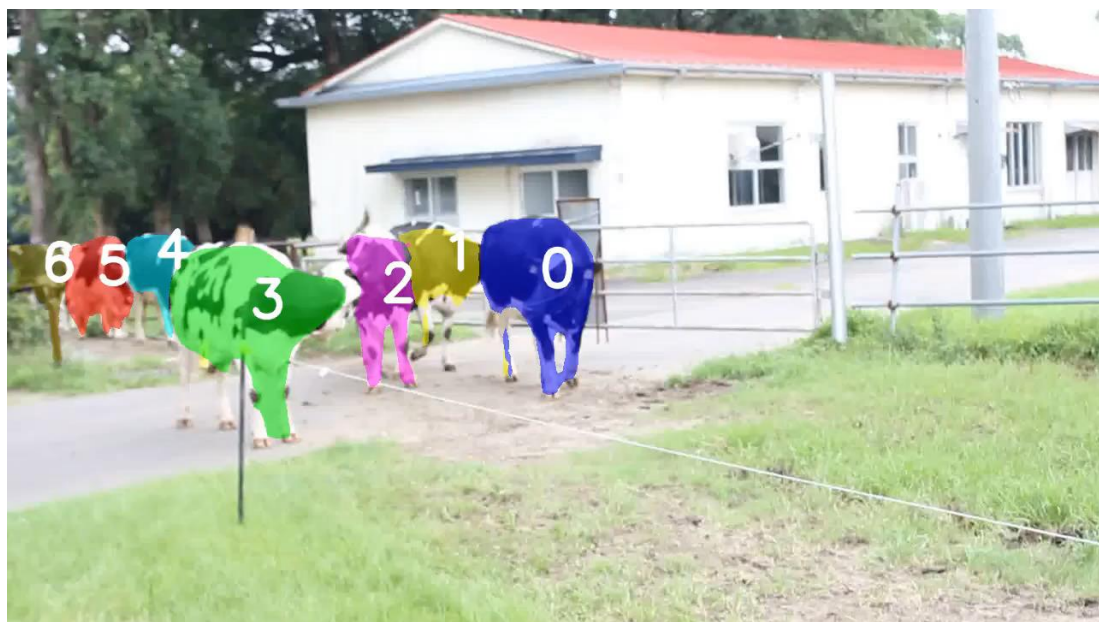
実験結果 (M724 : tr00)

提案手法のIoU値がほとんどのフレームにおいて0.9を超える高い精度



実験結果 (M724 : tr00)

提案手法のIoU値がほとんどのフレームにおいて
0.9を超える高い精度

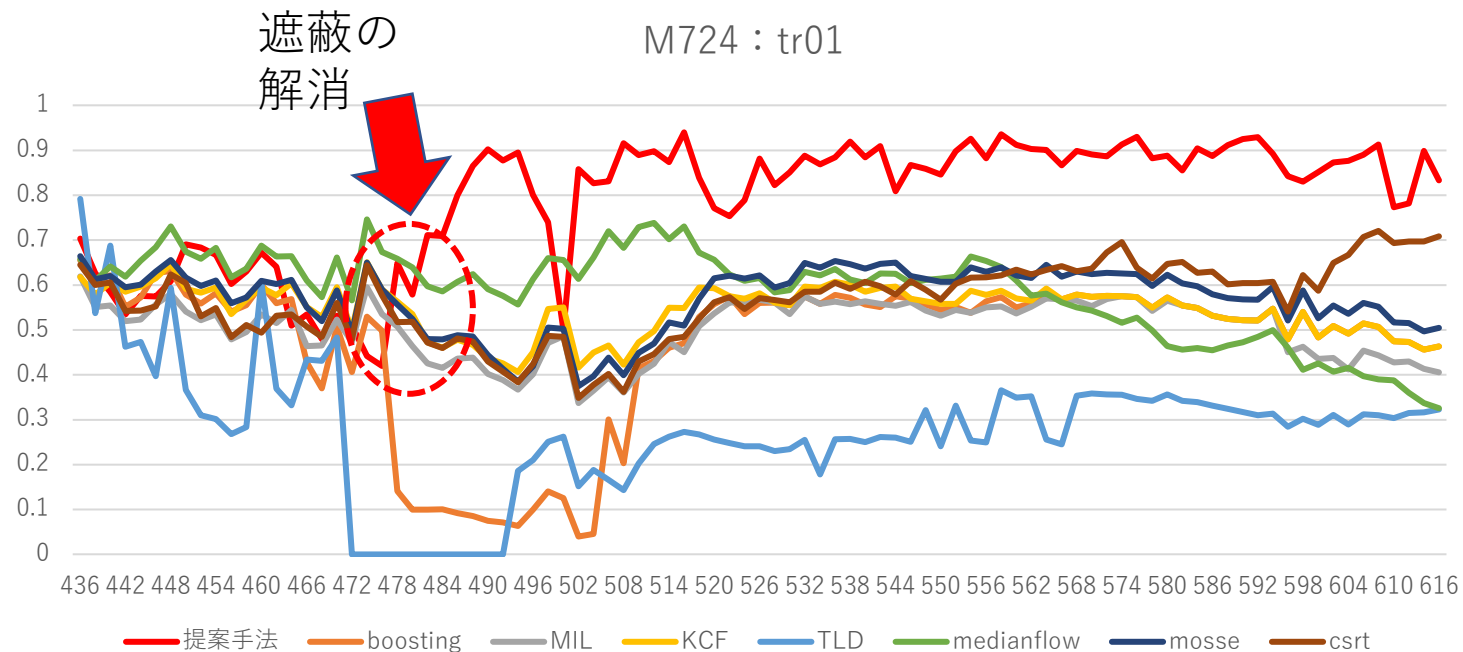


実験結果 (M724 : tr01)

470~490フレーム周辺に着目

従来手法：IoU値が低下

提案手法：IoU値を維持

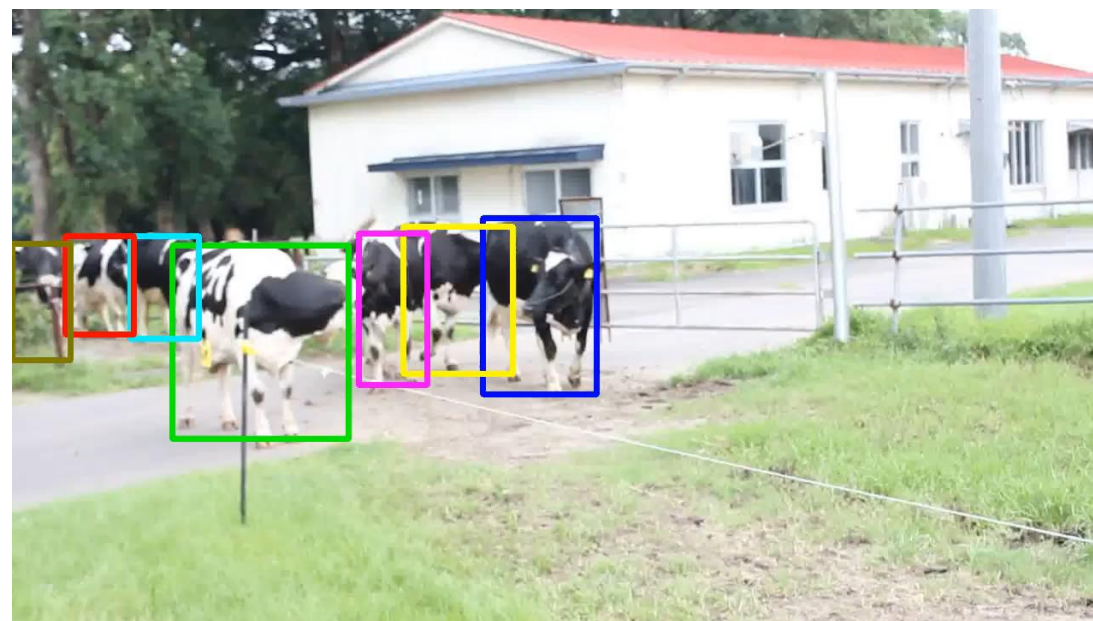


実験結果 (M724 : tr01)

470~490フレーム周辺に着目

従来手法 : IoU値が低下

提案手法 : IoU値を維持



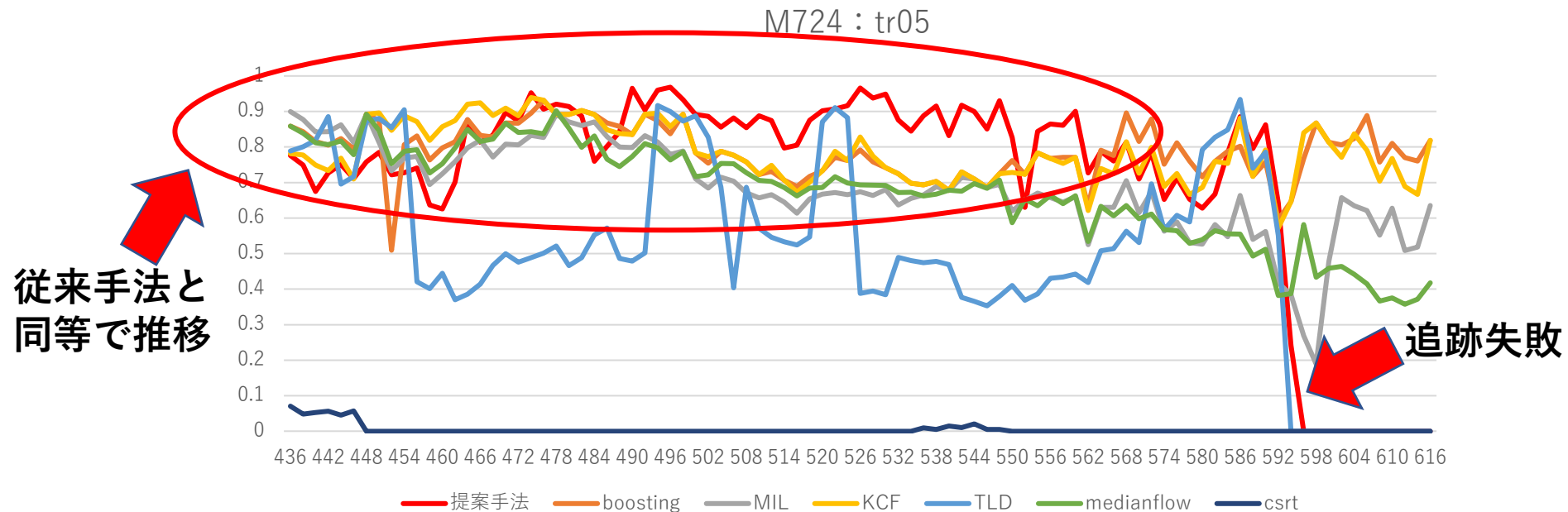
実験結果 (M724 : tr05)

592フレーム周辺までは従来手法と同等の精度

592フレーム周辺から他の物体を追跡 (IDスイッチ)

手ブレの発生 = 疑似的な高速の移動

→牛の移動が遅い = 数フレームで大きく移動しないという前提のため見失った



まとめ

牛の個体管理のために高精度な追跡の実現

- インスタンス・セグメンテーションを利用した高精度な個体検出
- 検出結果を時間方向に対応付けた追跡

比較実験により有効性を評価

提案手法は遮蔽による見え方の変化に対応して追跡を行える

今後の課題

より移動の速い類似個体同士の物体追跡

補足

- 以下
- Faster R-CNNとMask R-CNNの違い
- C123 (tr01) の実験結果
- M358の実験結果
- M724 (tr02,tr03,tr04,tr06) の実験結果

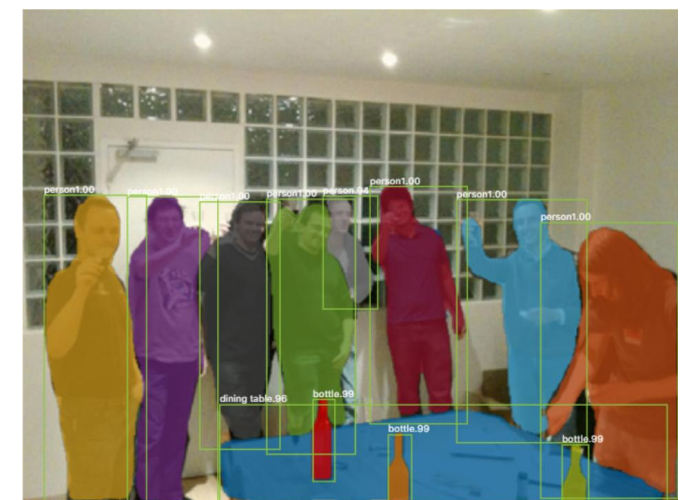
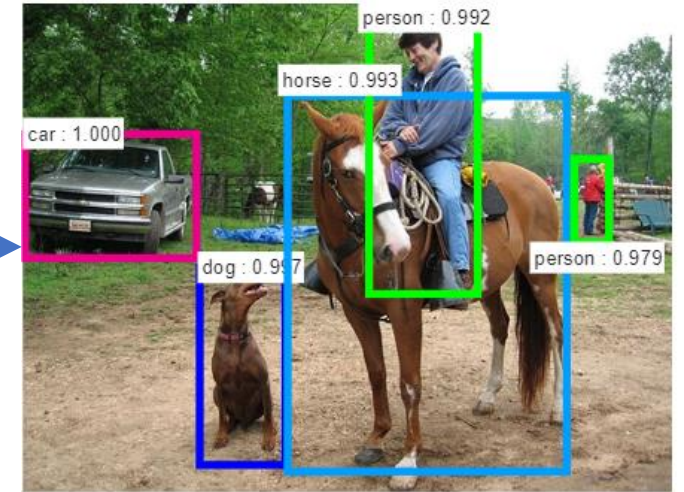
インスタンス・セグメンテーション

- 過去のディープラーニングを用いた物体検出との違い

- 過去：Faster R-CNN

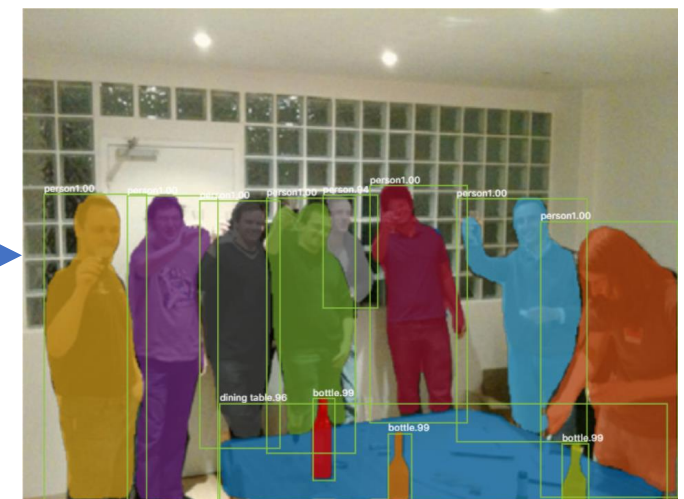
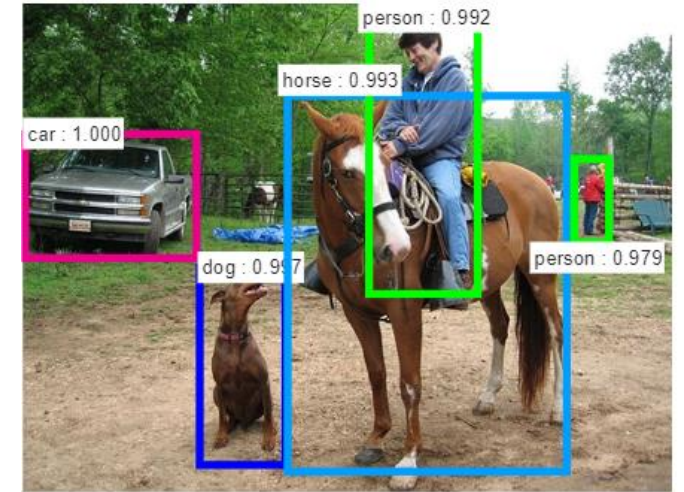
バウンディングボックスによる物体検出
+ 種類（クラス）の分類

→バウンディングボックスによる対象の検出は**対象以外**の情報が多い



インスタンス・セグメンテーション

- 過去のディープラーニングを用いた物体検出との違い
- 使用する手法：Mask R-CNN
 - バウンディングボックスによる物体検出
 - + 検出された領域内の各ピクセルのクラス
 - + 種類（クラス）の分類
 - より正確な検出結果を用いて追跡



Faster R-CNN

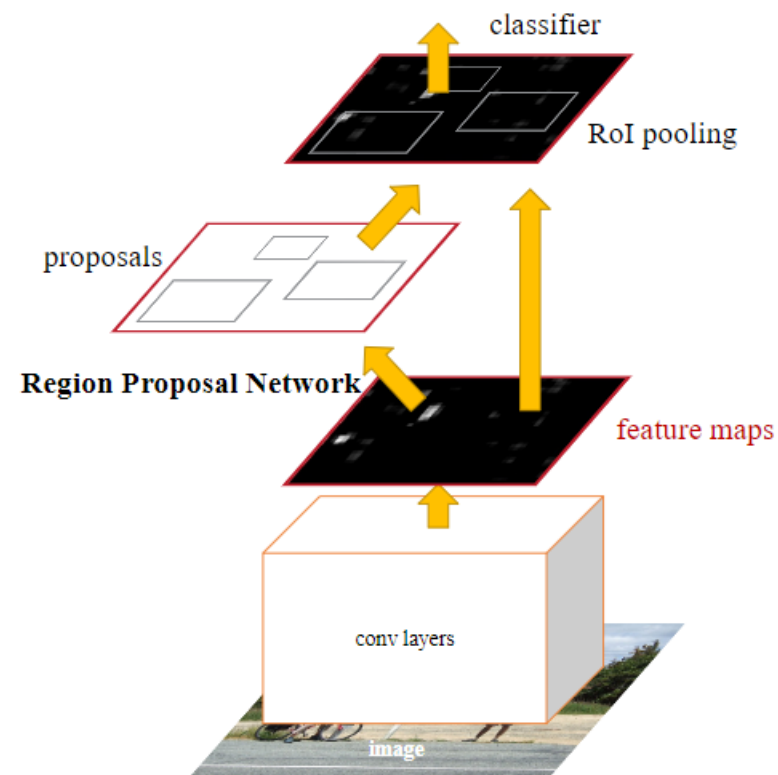
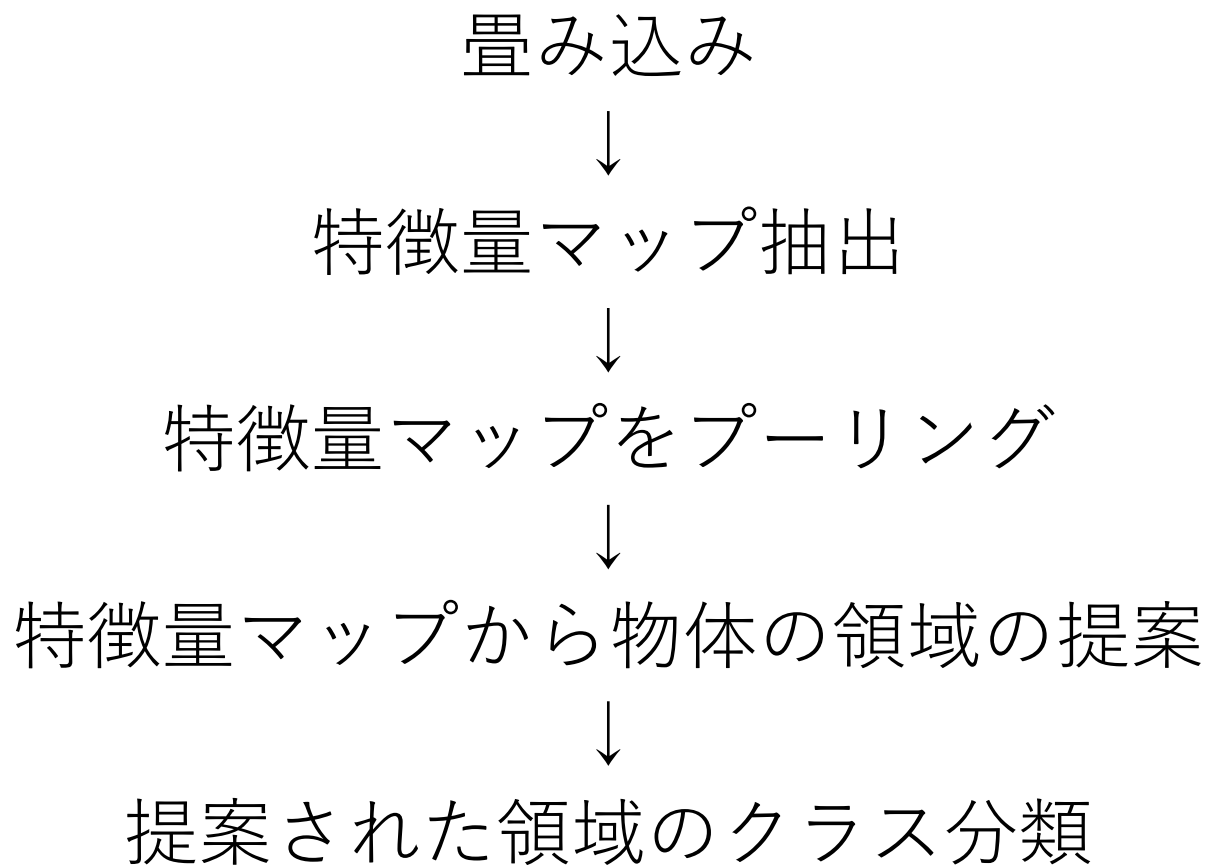


Figure 2: Faster R-CNN is a single, unified network for object detection. The RPN module serves as the 'attention' of this unified network.

Mask R-CNN

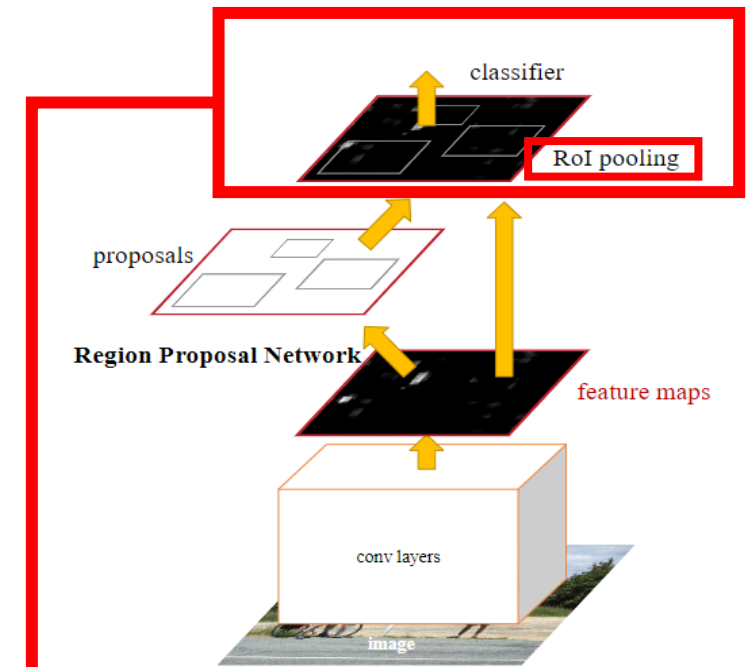
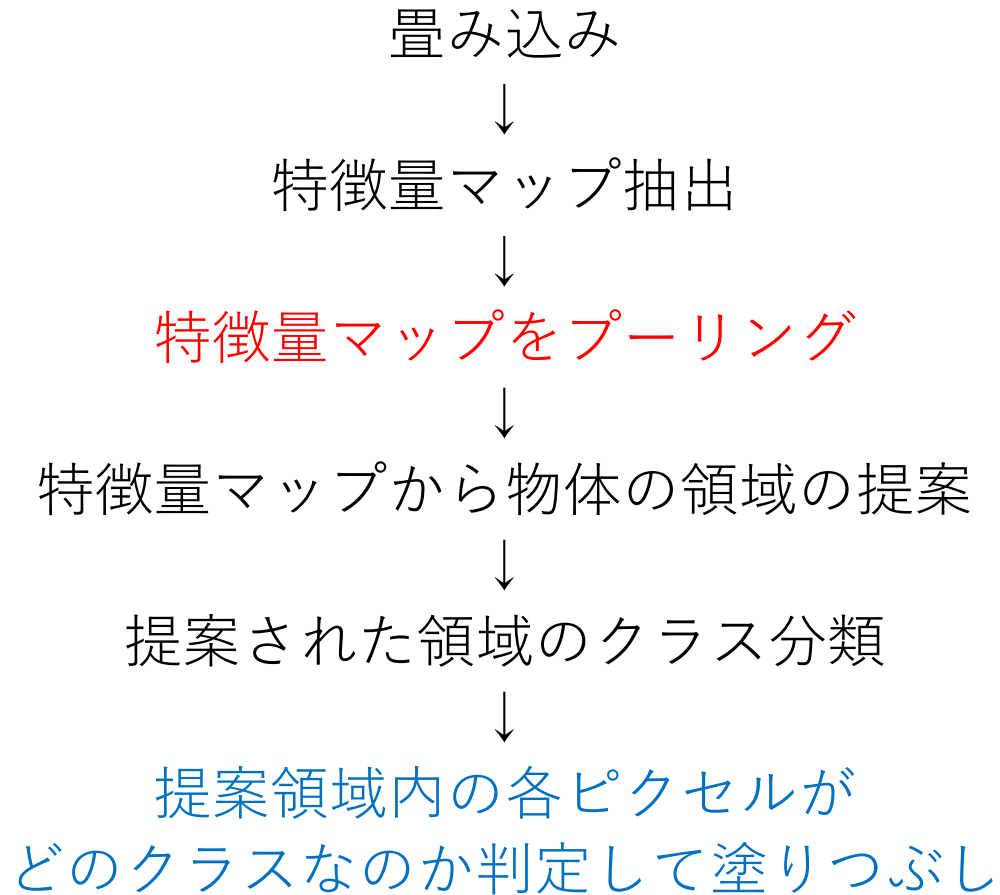


Figure 2: Faster R-CNN is a single, unified network for object detection. The RPN module serves as the 'attention' of this unified network.

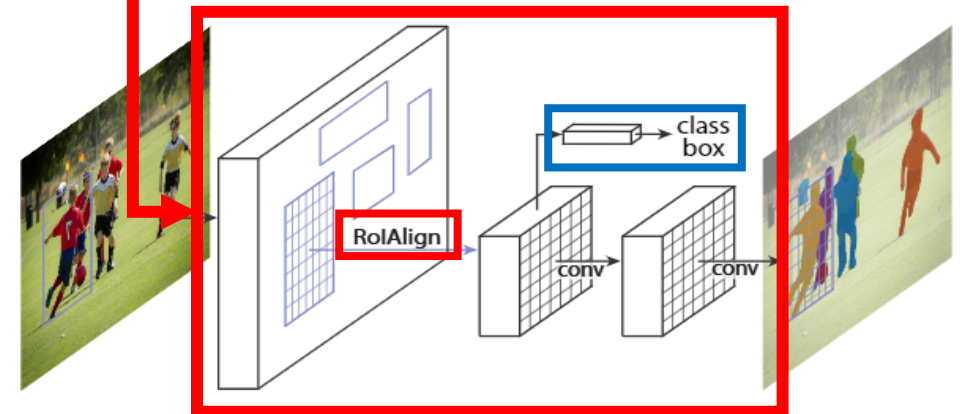


Figure 1. The Mask R-CNN framework for instance segmentation.

Mask R-CNN

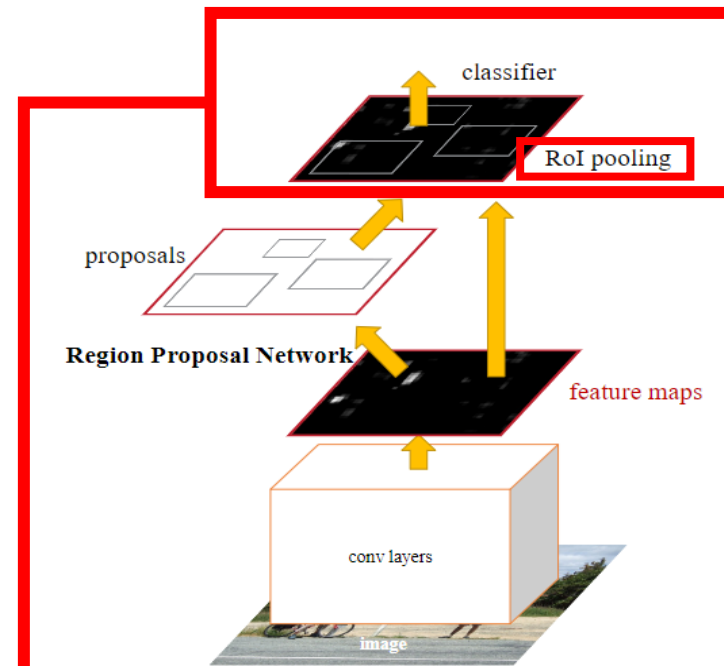
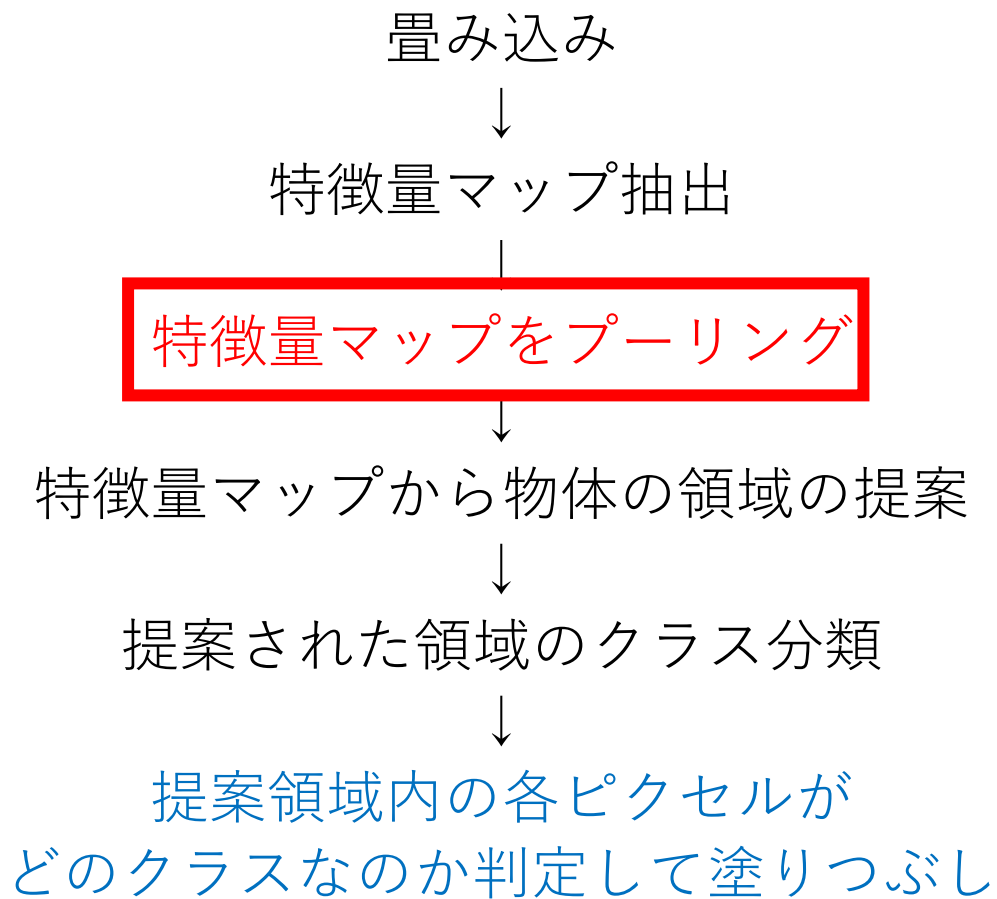


Figure 2: Faster R-CNN is a single, unified network for object detection. The RPN module serves as the 'attention' of this unified network.

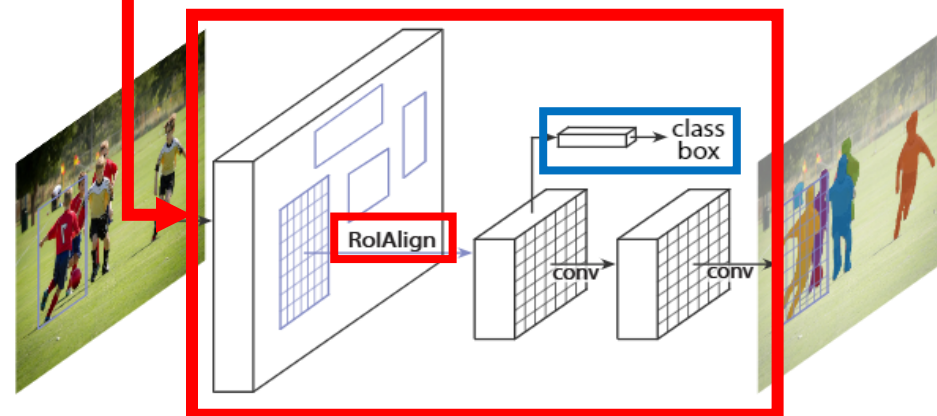


Figure 1. The Mask R-CNN framework for instance segmentation.

Mask R-CNN

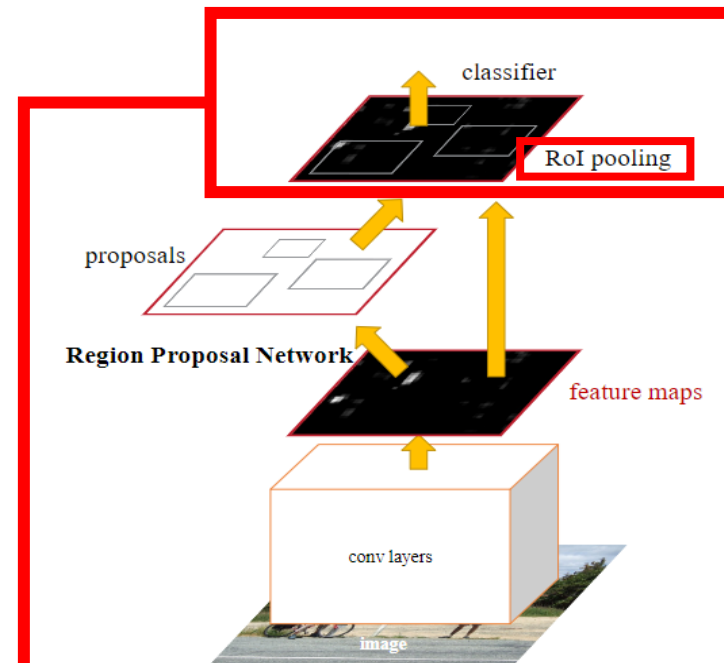
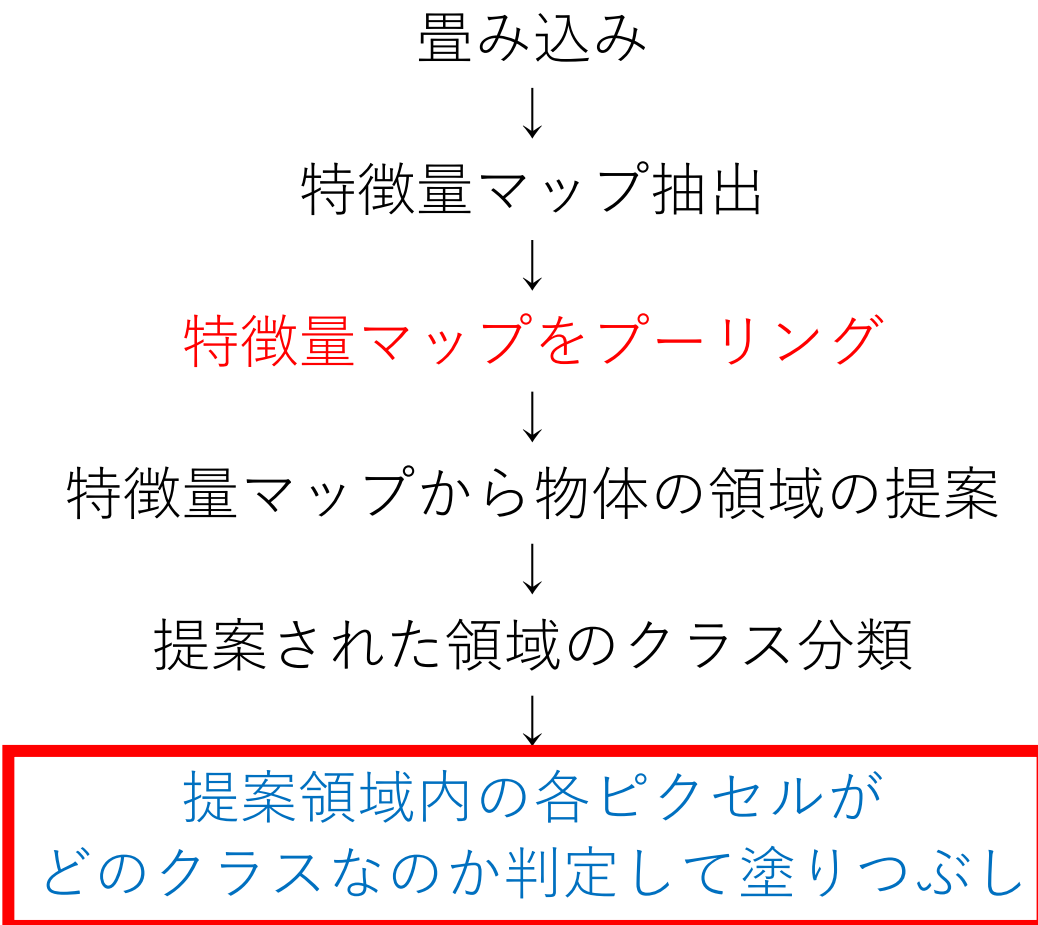


Figure 2: Faster R-CNN is a single, unified network for object detection. The RPN module serves as the 'attention' of this unified network.

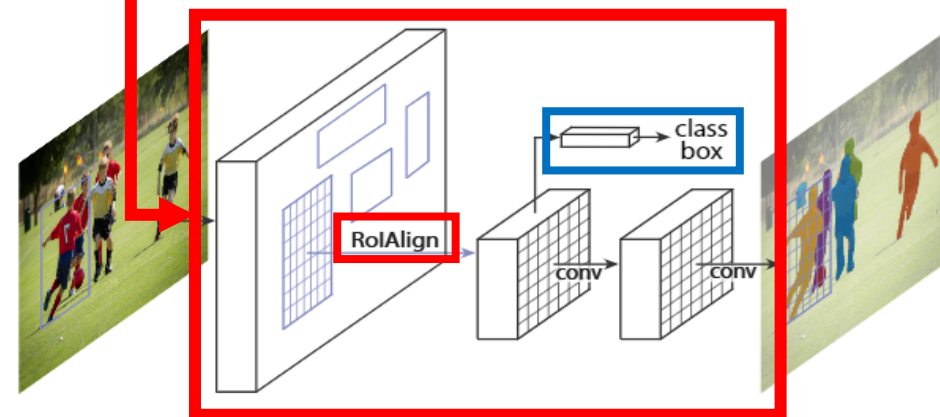


Figure 1. The Mask R-CNN framework for instance segmentation.

C123：画像例

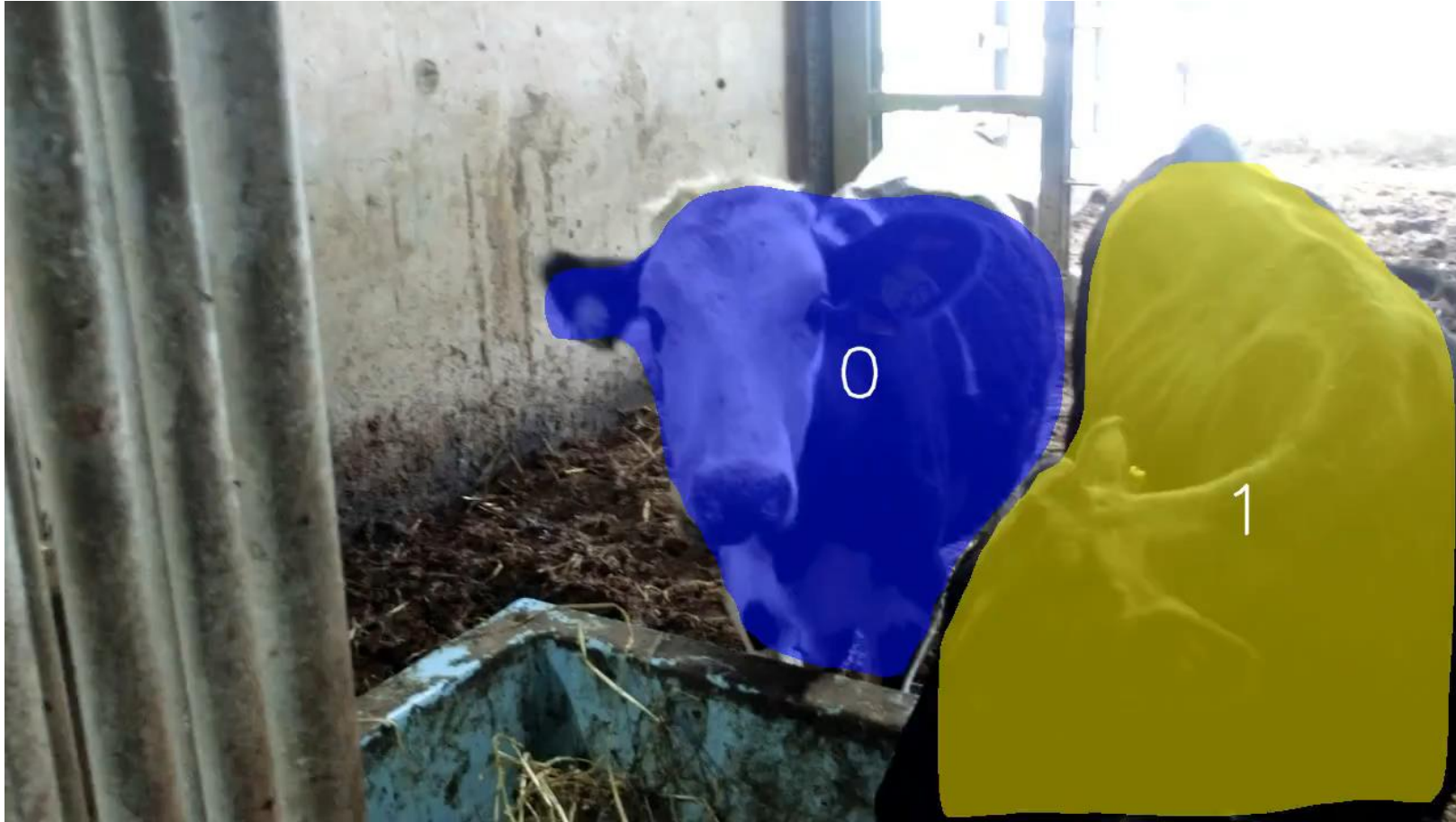


C123：正解データ

- マスク領域で表す



実験結果 (C123)

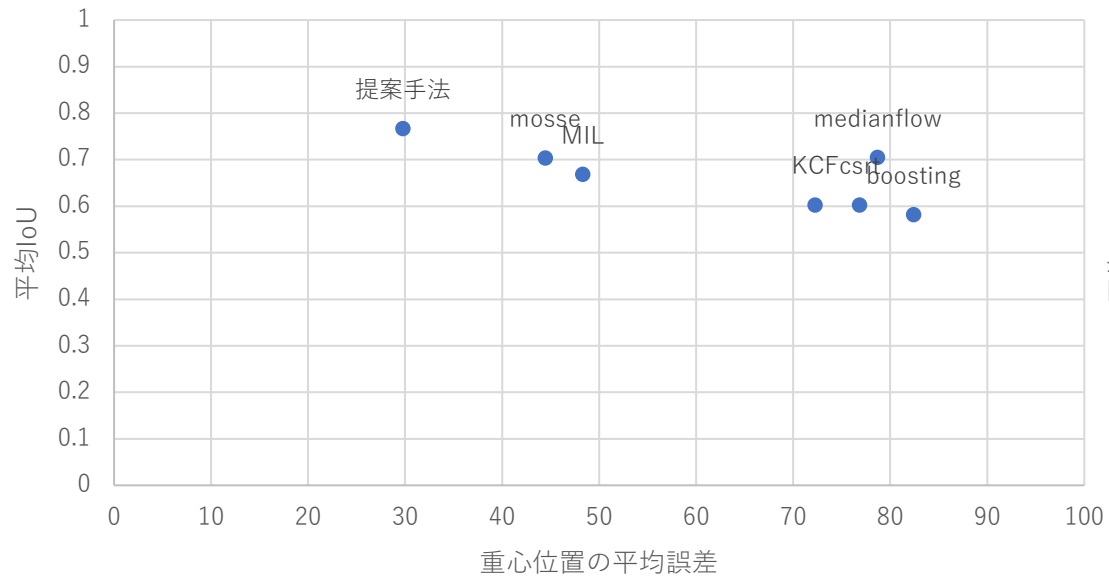


実験結果 (C123)

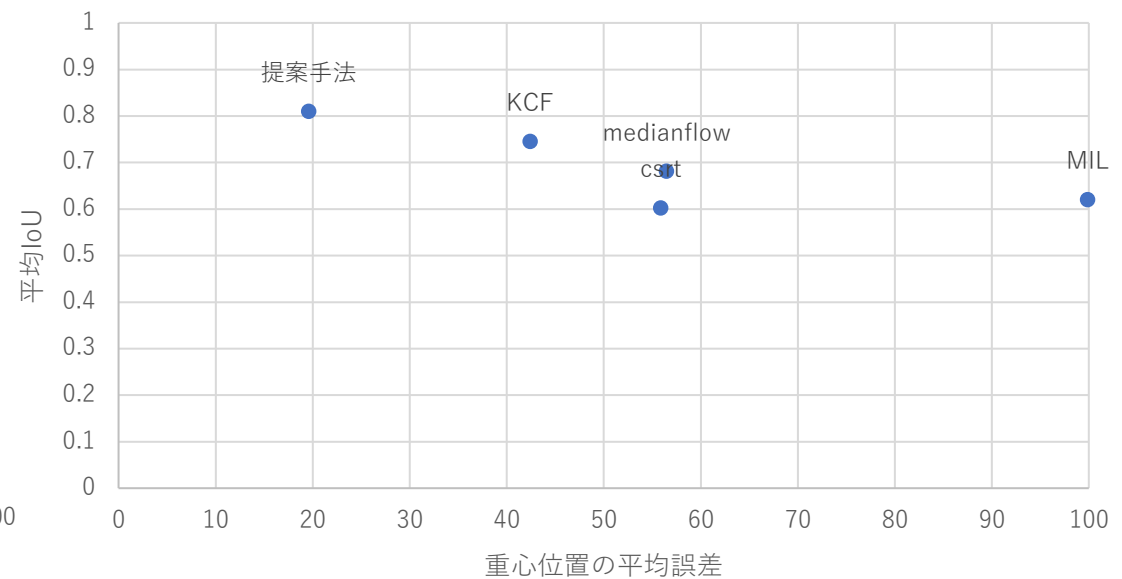
縦軸：全フレームの平均IoU

横軸：正解データと追跡領域の重心位置の平均誤差

C123 : tr00 (左)



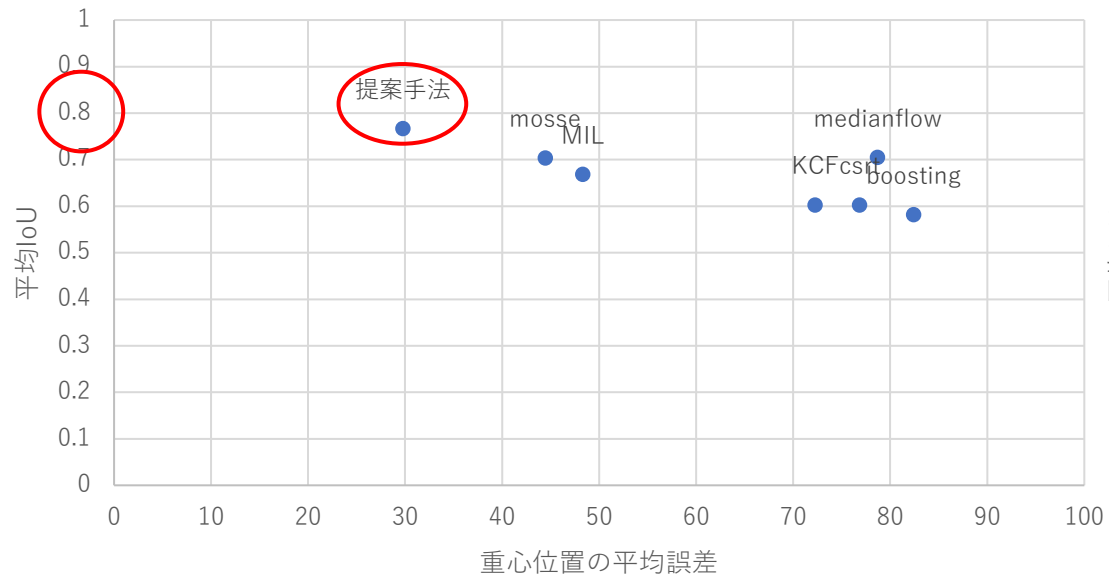
C123 : tr01 (右)



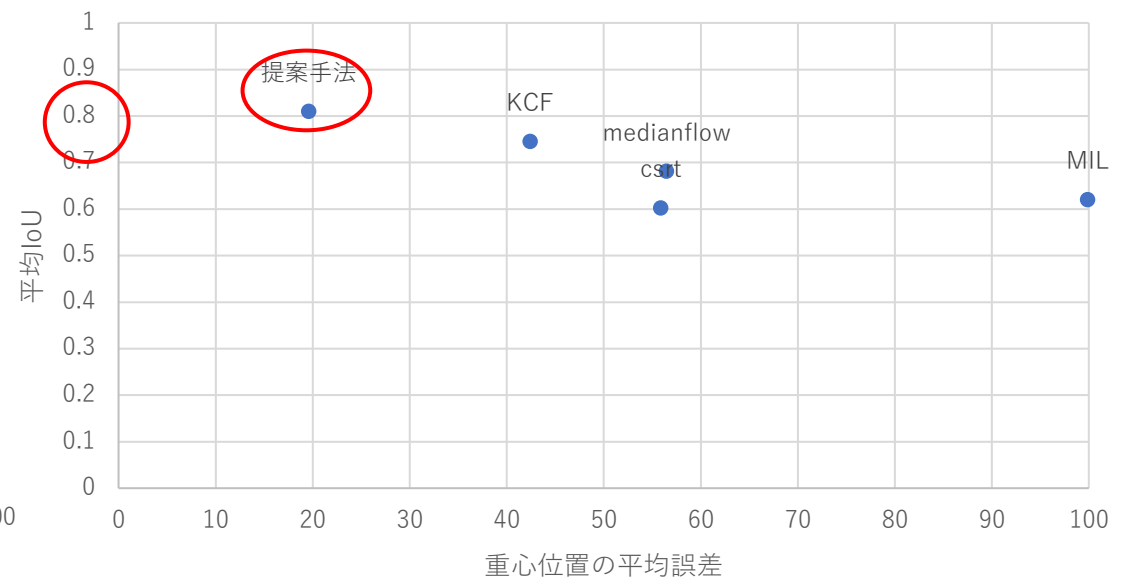
実験結果 (C123)

両対象に対し追跡手法として高い性能を示した
従来手法より平均IoU、平均誤差の精度が良い
→提案手法の有効性を示した

C123 : tr00 (左)

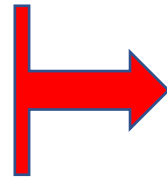


C123 : tr01 (右)



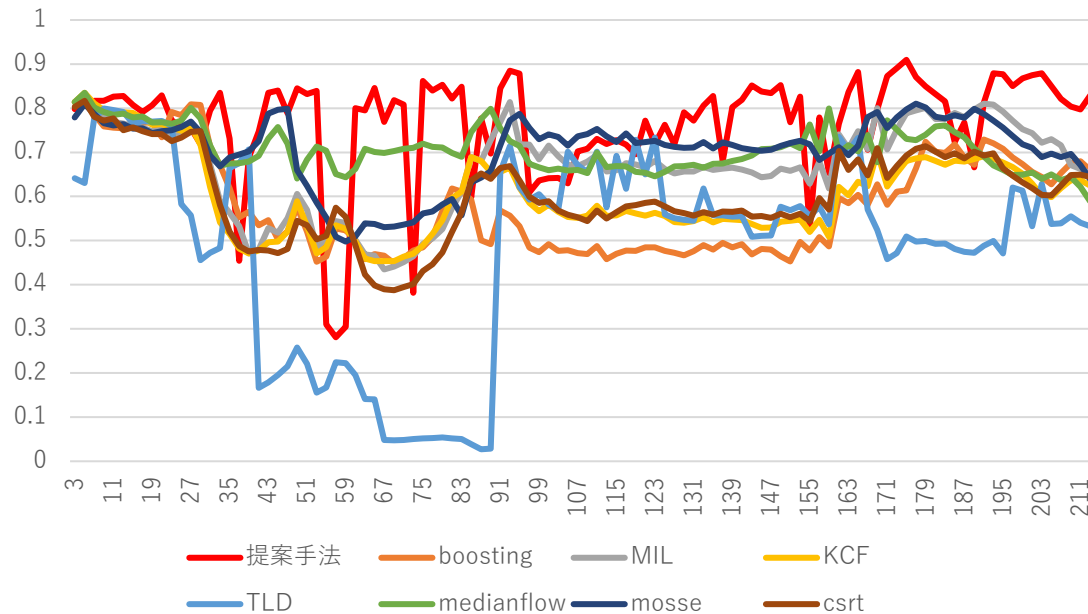
実験結果 (C123:tr00左)

横軸：フレーム
縦軸：IoU値



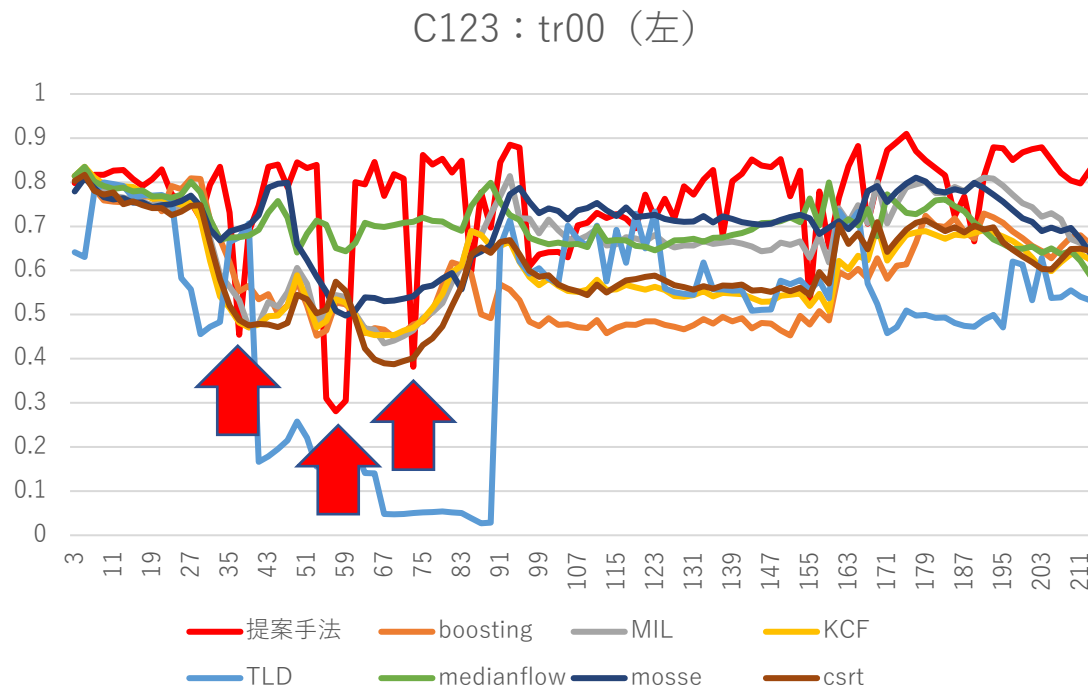
フレームの小さい順にからIoU値の推移を見ることで追跡結果の確認

C123 : tr00 (左)



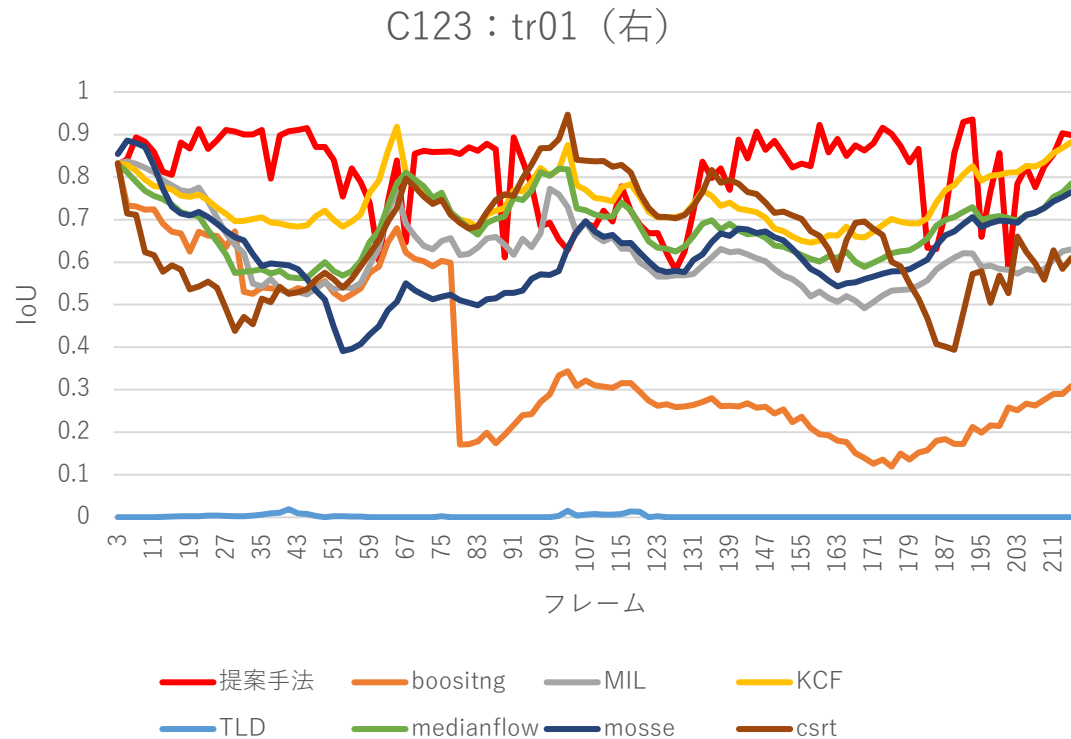
実験結果 (C123:tr00左)

局所的に検出できていないフレームが存在するときにIoU値が低く出ている場合がある



実験結果 (C123:tr01)

局所的に検出できていないフレームが存在するときにIoU値が低く出ている場合がある



実験結果 (M358)

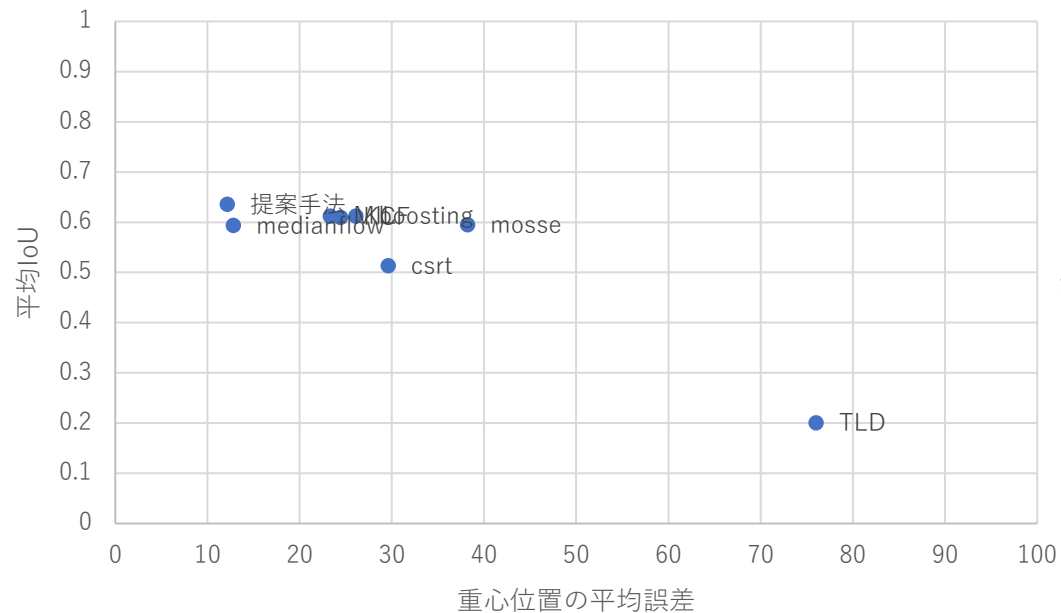


実験結果 (M358)

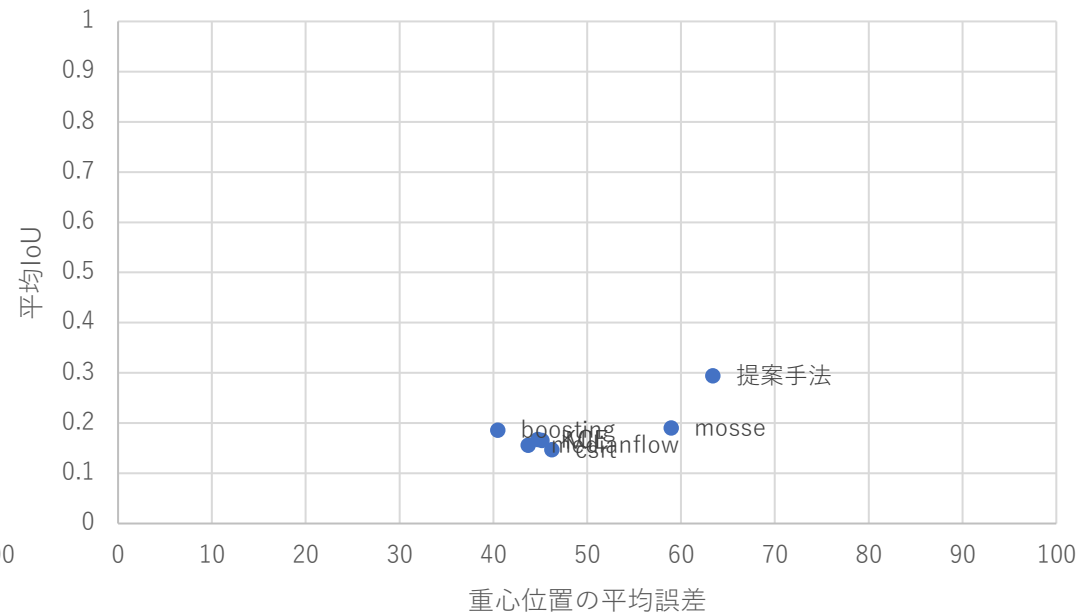
縦軸が全フレームの平均IoU

横軸が正解データの重心距離との距離の平均誤差

M358 : tr00 (手前)



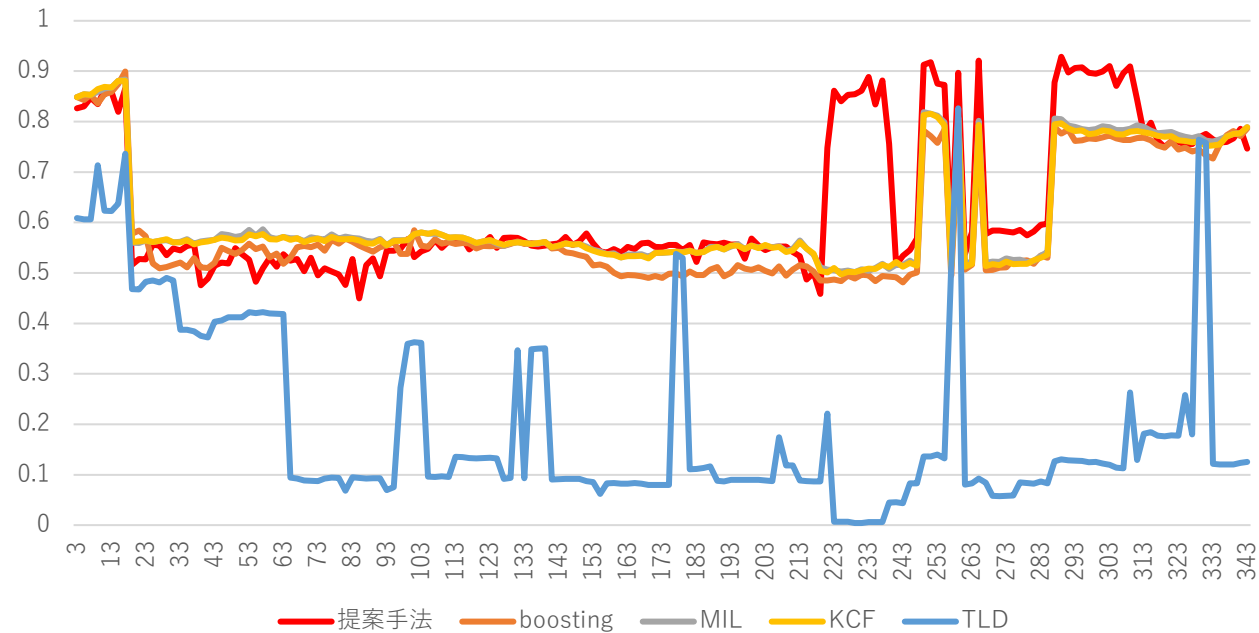
M358 : tr01 (後方)



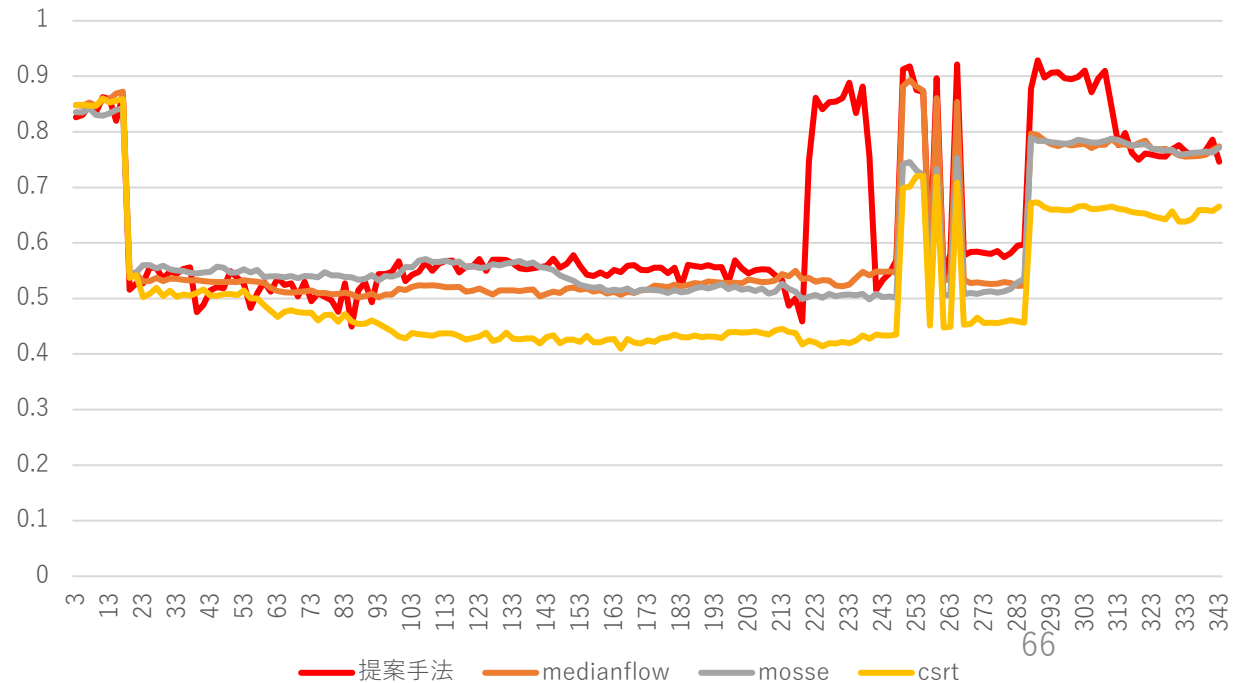
実験結果 (M358:手前)

従来手法と同等の結果が多いが検出がうまくいっている場合が見られる

M358:手前

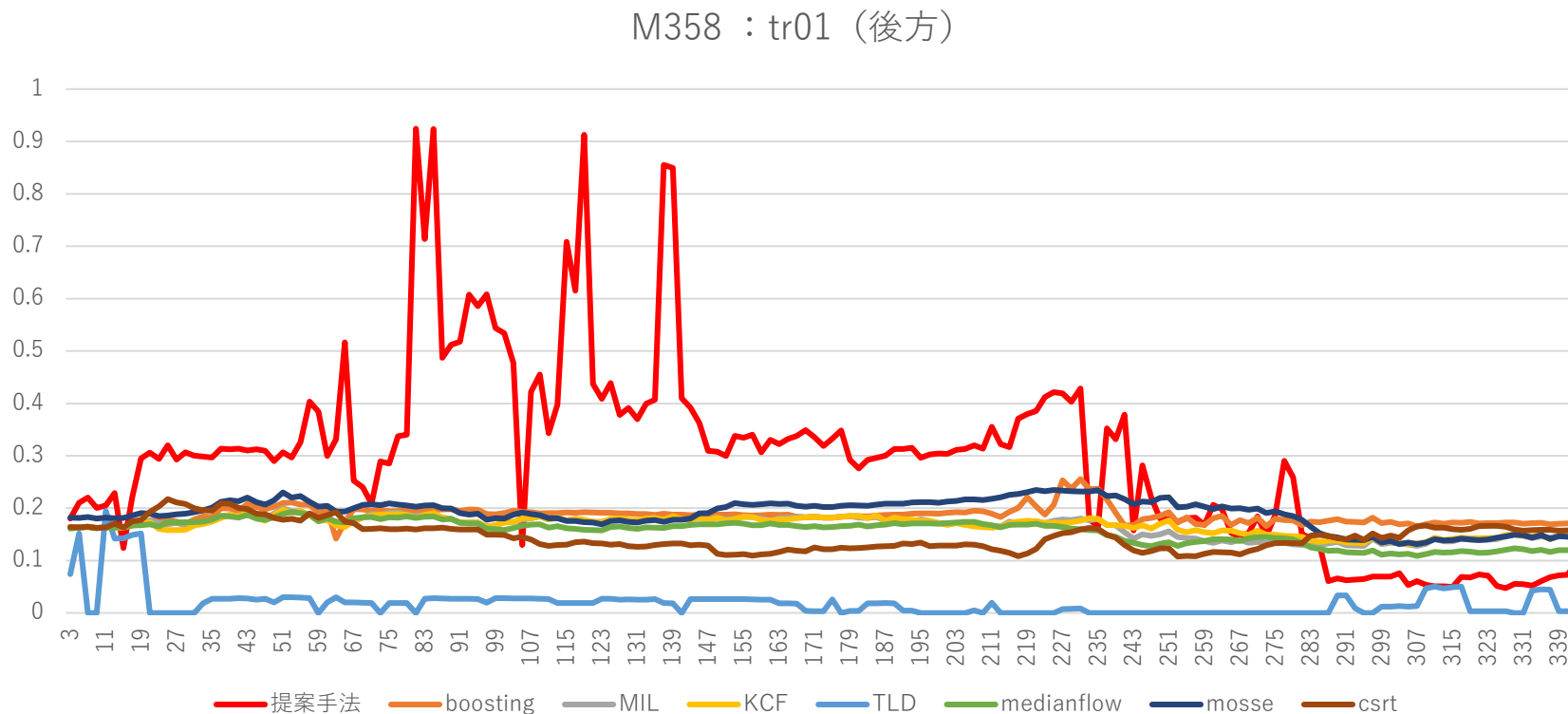


M358:手前



実験結果 (C358:tr01後方)

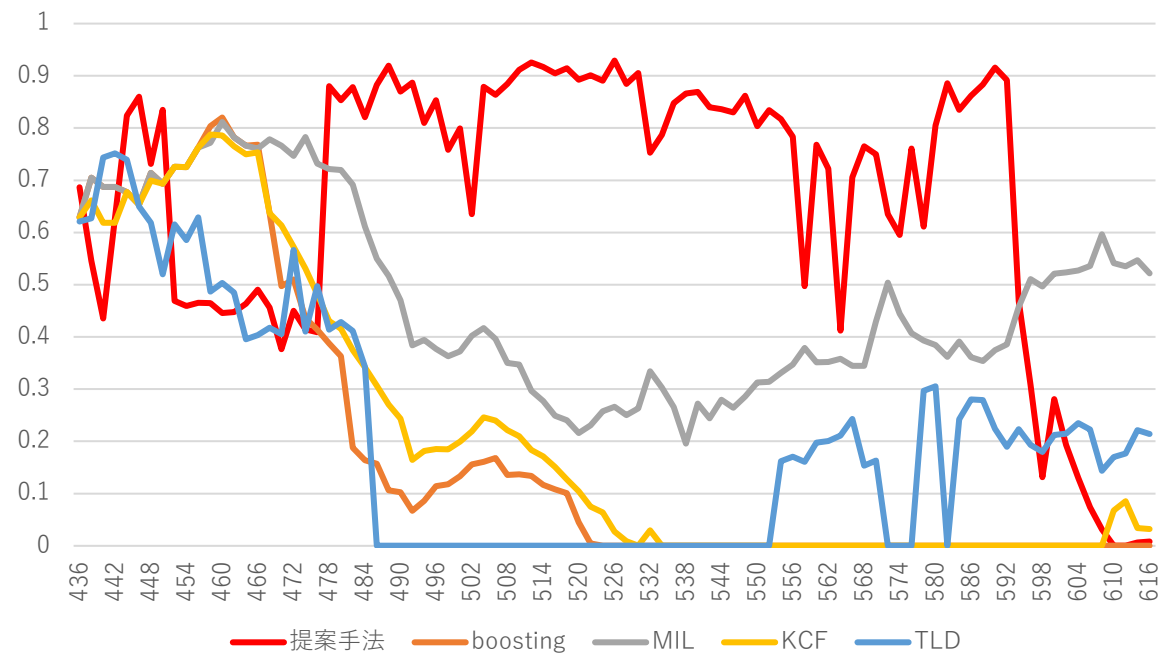
局所的に検出できていないフレームが存在するときにIoU値が低く出ている場合がある



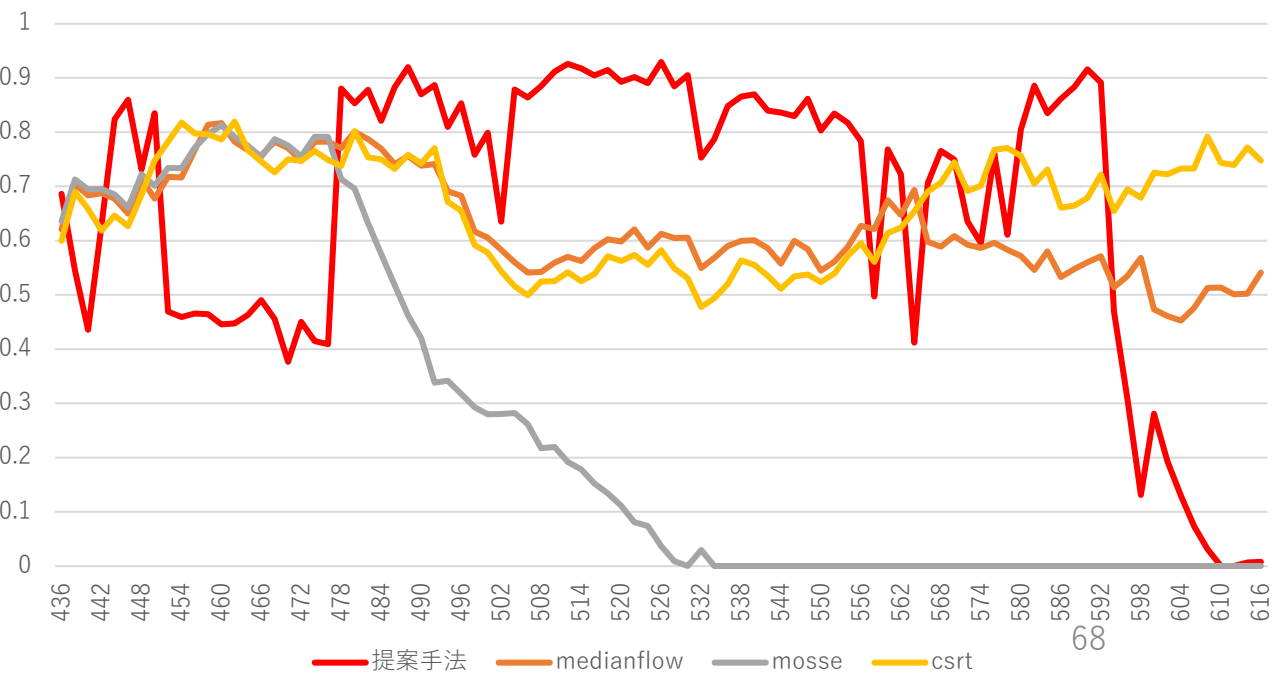
実験結果 (M724 : tr02)

検出結果

M724 : tr02



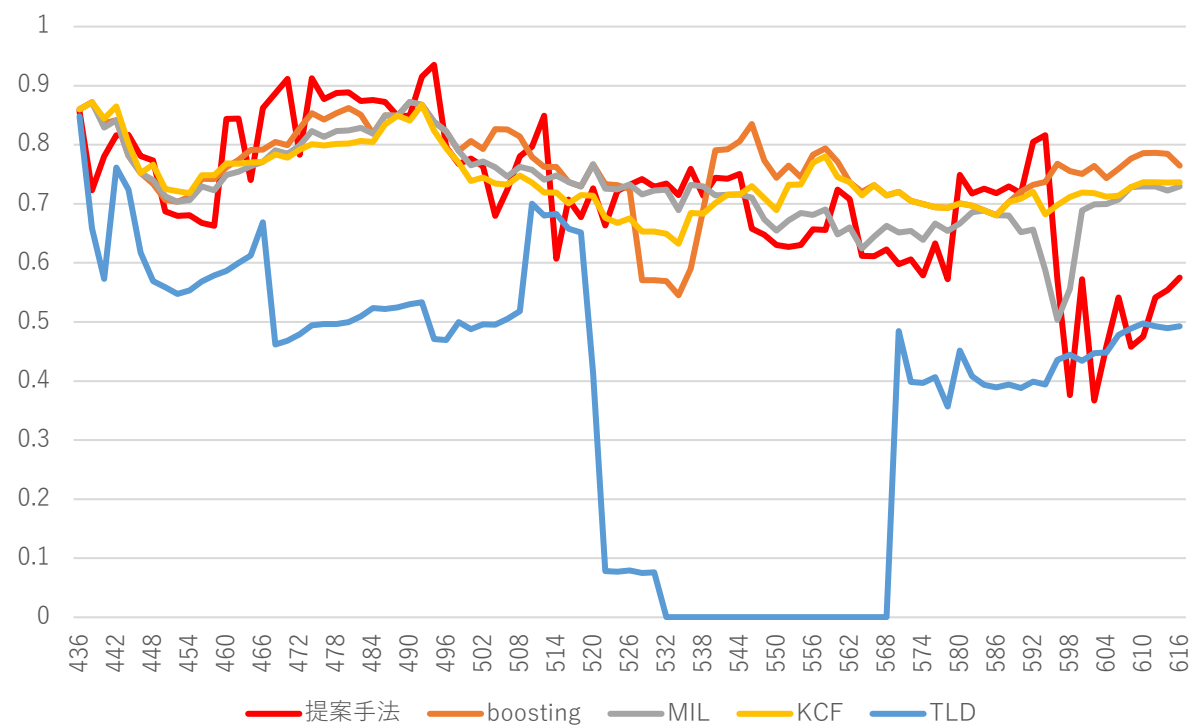
M724 : tr02



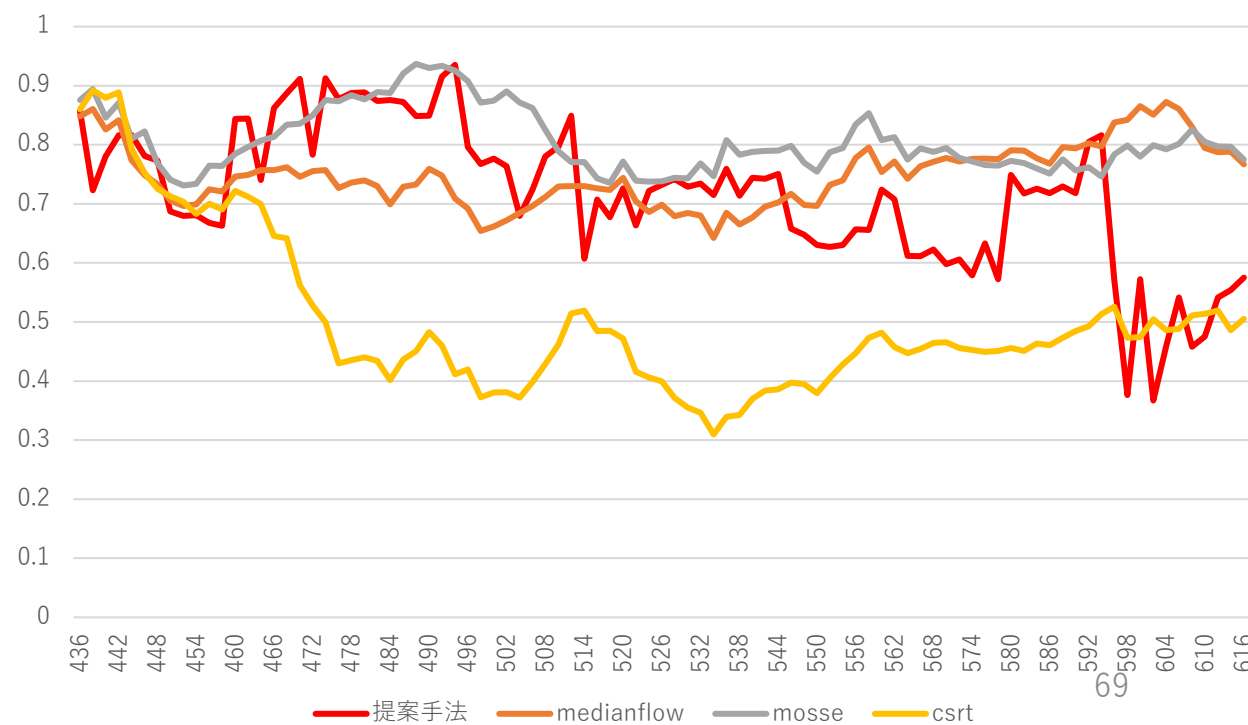
実験結果 (M724 : tr03)

検出結果

M724 : tr03



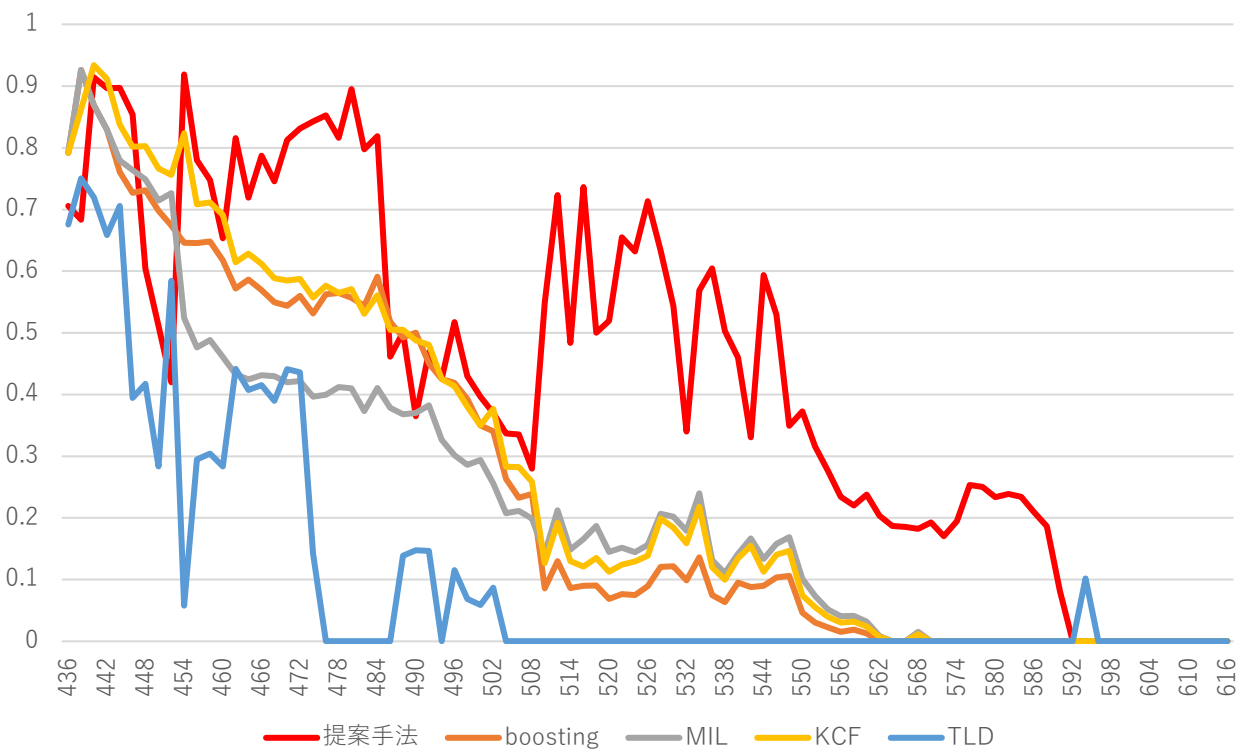
M724 : tr03



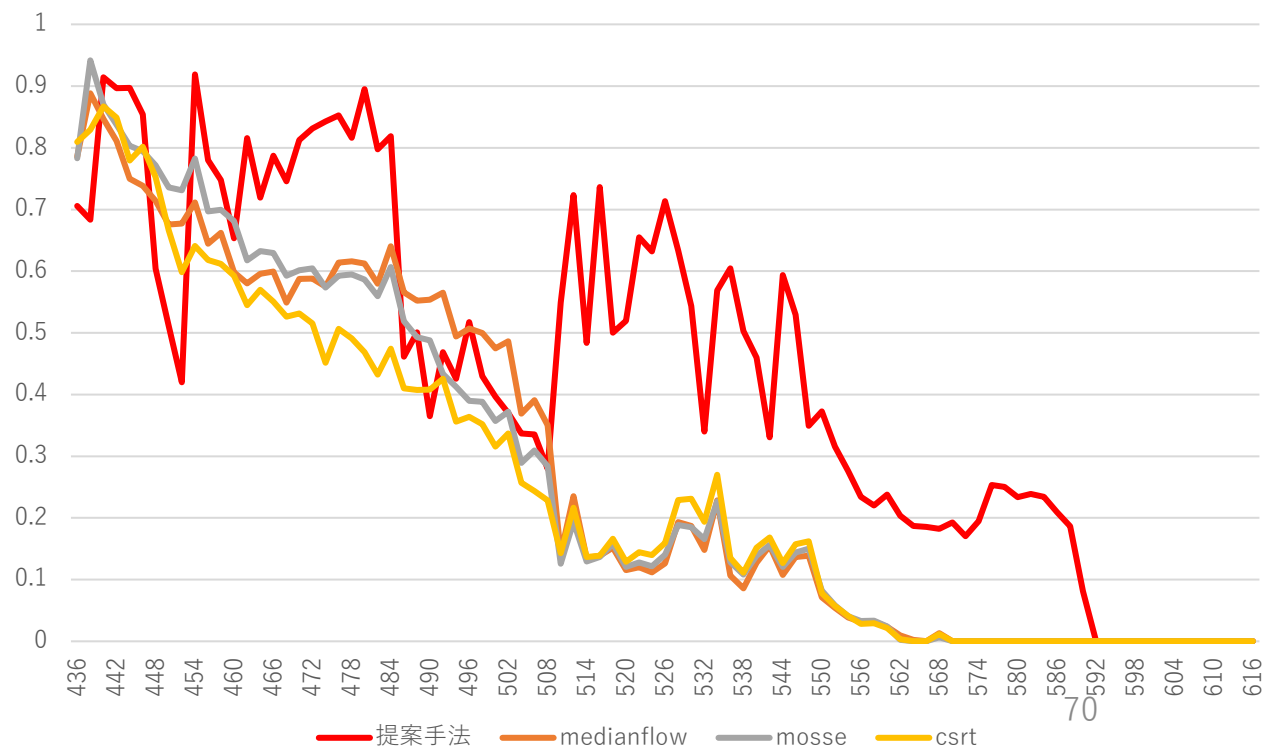
実験結果 (M724 : tr04)

検出結果

M724 : tr04

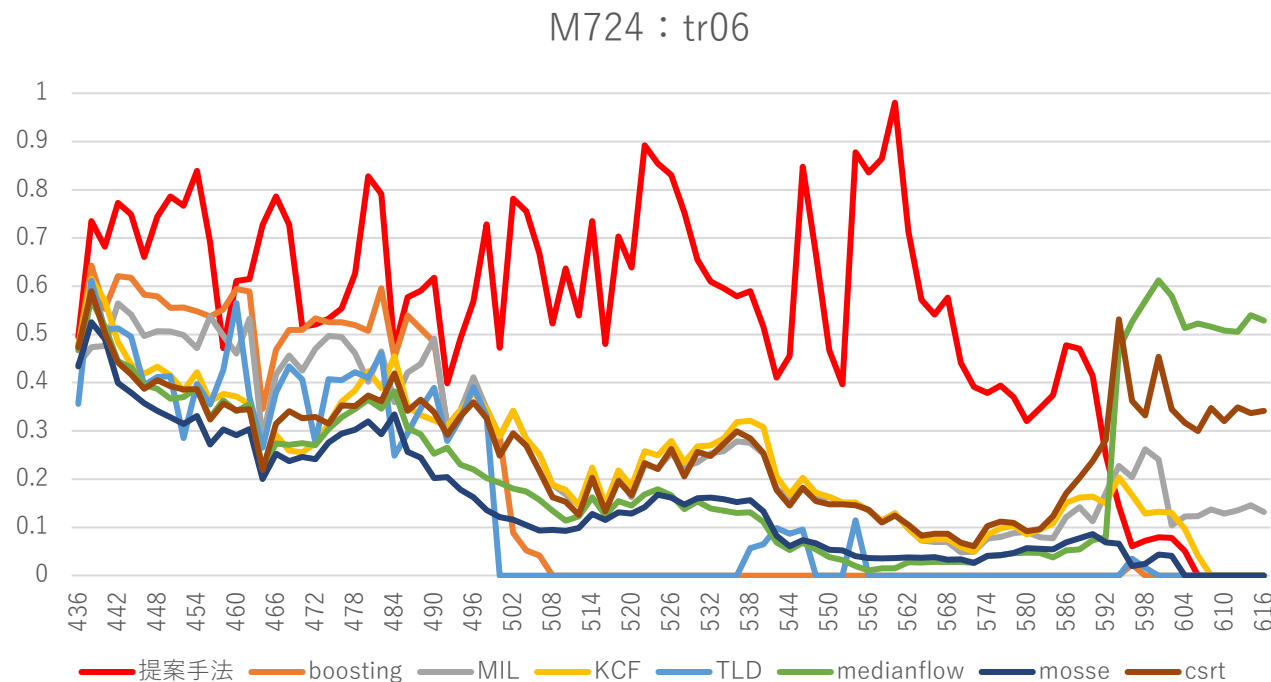


M724 : tr04



実験結果 (M724 : tr06)

tr05と同様に592フレーム周辺からIoU値が下がった
→tr05の追跡が手ブレによりtr06を追跡対象と間違えたため



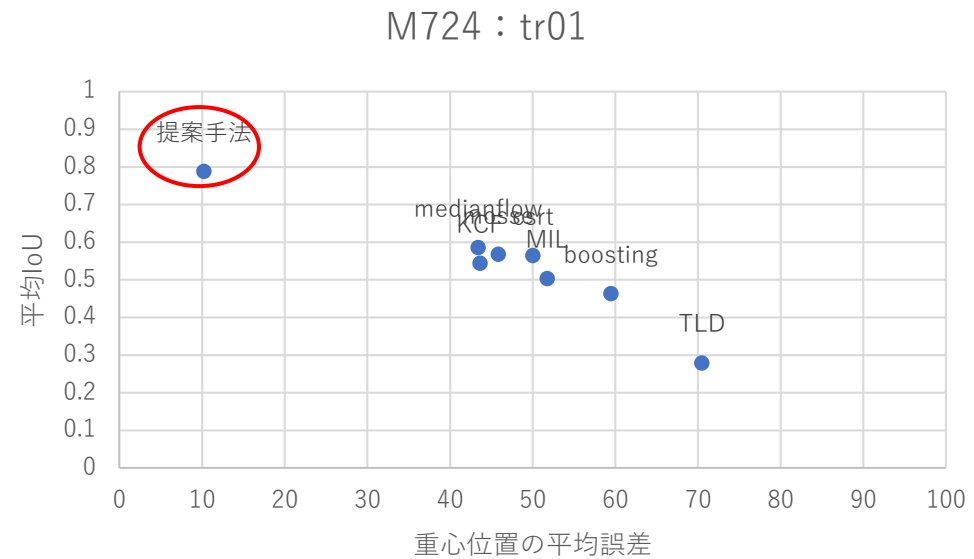
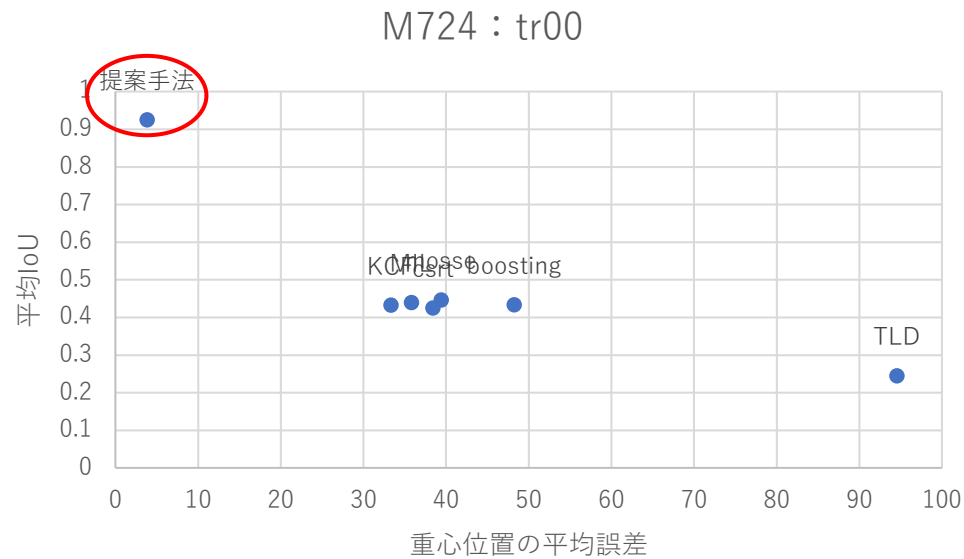
課題

- 提案手法とバウンディングボックスを利用した場合との追跡の比較
 - 提案手法はバウンディングボックスを利用した場合を上回る精度か
- マスクレベルでの追跡の評価
 - より正確な領域での追跡で精度向上
- より移動の速い類似個体同士の物体追跡
 - 領域の重なり以外での評価基準の導入等の改良

実験結果 (M724)

良好な結果：tr00, tr01, tr04, tr06

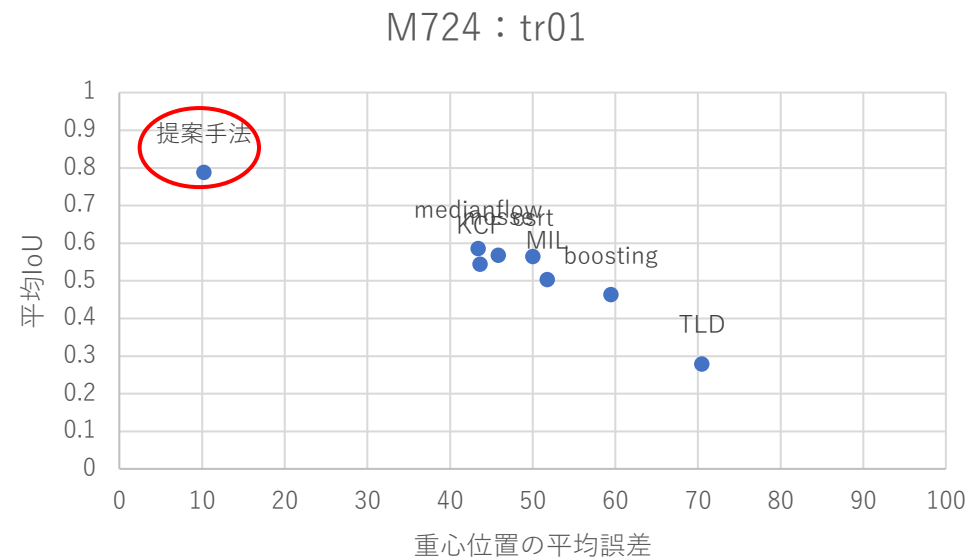
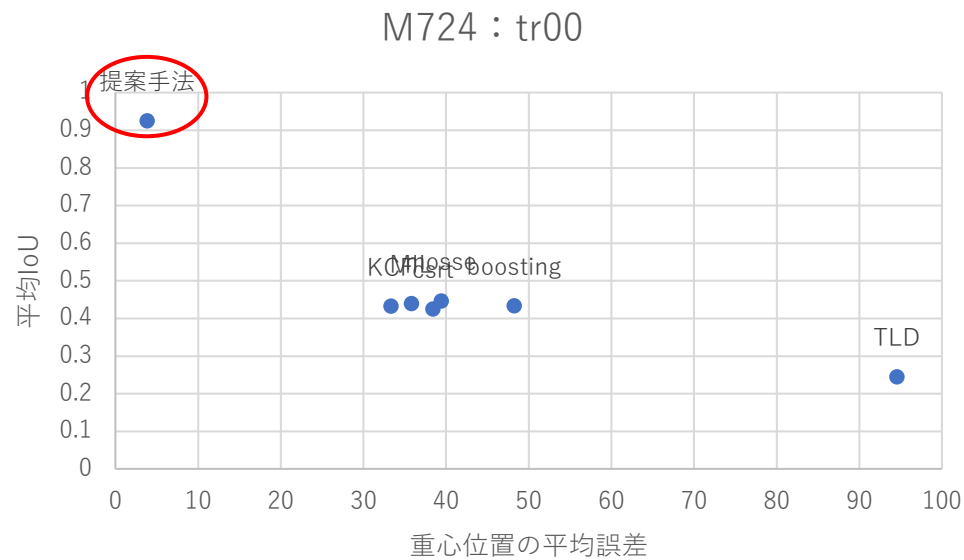
従来手法と同等：tr02, tr03, tr05



実験結果 (M724)

良好な結果：tr00, tr01, tr04, tr06

従来手法と同等：tr02, tr03, tr05

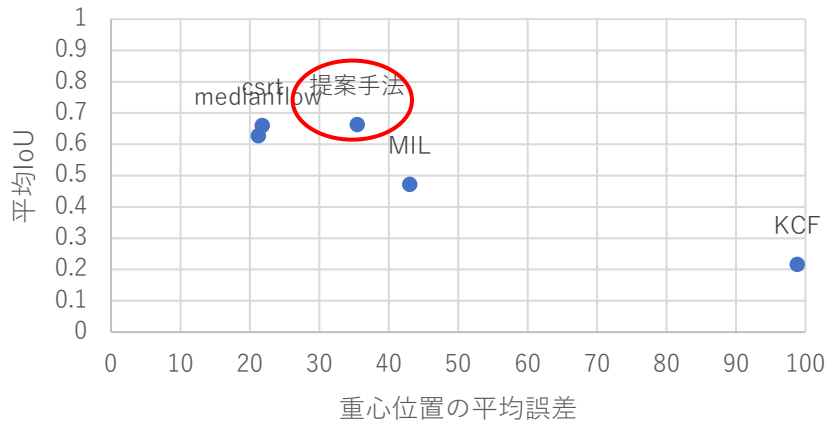


実験結果 (M724)

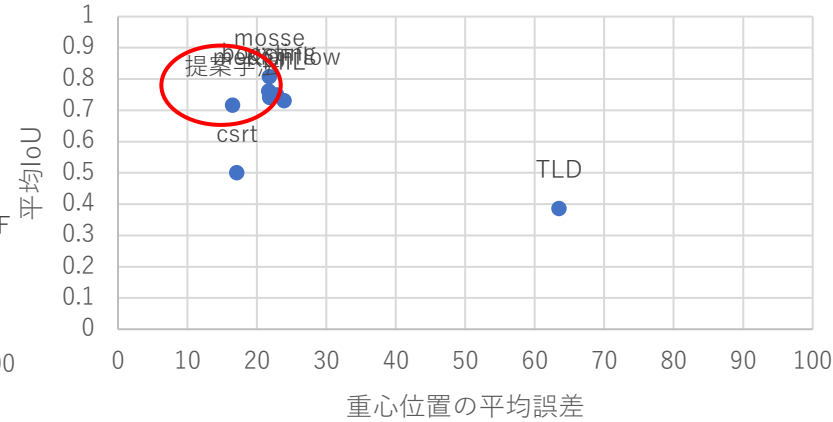
良好な結果：tr00, tr01, tr04, tr06

従来手法と同等：tr02, tr03, tr05

M724 : tr02



M724 : tr03



M724 : tr05

