

## 1. はじめに

多品種ばら積みのピッキングにおいて、オクルージョンが生じている物体を把持する場合、他物体が上に存在するため、把持に失敗するという問題がある。この問題を解決するには、上に存在する物体から把持していく必要があり、多品種ばら積みのピッキングでは物体間の上下関係を把握して把持戦略を決定する必要がある。この把持戦略を行うためには、物体の検出及び認識、物体の可視領域のセグメンテーション、物体のオクルージョン領域のセグメンテーションの3つのタスクを行わなければならない。

本研究では、これらを1つのモデルで実現する Occluded Mask R-CNN(OM R-CNN) および Interpolated Mask R-CNN(IM R-CNN) を提案する。また、推定結果をもとに物体間の上下関係を記述するツリー表現法を提案する。提案するネットワークモデルはインスタンスセグメンテーション手法である Mask R-CNN[1] をベースとし、物体のオクルージョン領域のセグメンテーションを行うブランチを導入して実現する。また、提案手法の推定結果をもとに物体の上下関係をツリー構造で構築することでシーンに対する最適な把持順序の構築ができる。

## 2. 提案手法

物体間の上下関係を把握して把持戦略を決定するには、物体の重なりを把握する必要がある。そのため、本研究では、シーン中全物体のオクルージョン領域推定手法および推定結果をもとに構築する物体間上下関係のツリー表現を提案する。

### 2.1 オクルージョン領域推定

物体間の重なりを把握するには、物体のオクルージョン領域を推定する必要がある。また、様々な大きさや形状のオクルージョンが生じている物体が存在するため、インスタンスごとにピクセルレベルでオクルージョン領域のセグメンテーションを行う Occluded Mask R-CNN(OM R-CNN) および Interpolated Mask R-CNN(IM R-CNN) を提案する。提案手法のネットワーク構成を図1に示す。

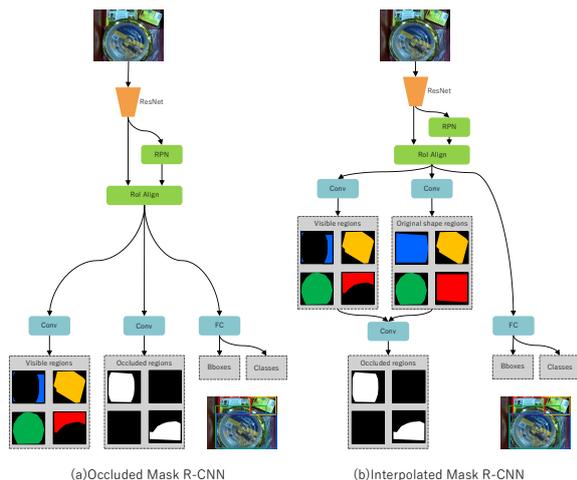


図1：提案手法のネットワーク構成

### Occluded Mask R-CNN

OM R-CNN は、インスタンスセグメンテーション手法である Mask R-CNN をベースとすることで、物体検出および認識、各インスタンスに対する可視領域のセグメンテーションを同時に行うことが可能である。図1(a)に示すように、Mask R-CNN に物体のオクルージョン領域のセグメンテーションタスクを追加することでオクルージョ

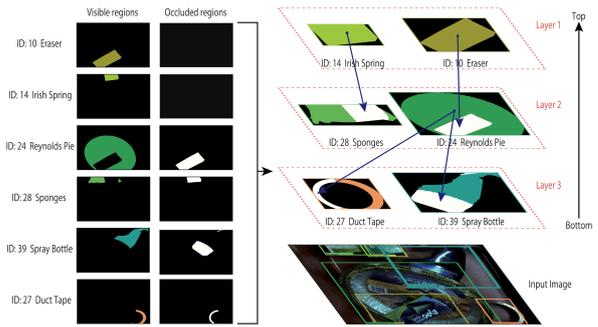


図2：上下関係のツリー表現の例

ン領域のセグメンテーションタスクを可能とする。提案手法では、物体検出、クラス分類、可視領域のセグメンテーション、オクルージョン領域のセグメンテーションのすべてを同時に学習する。

### Interpolated Mask R-CNN

オクルージョン領域は隠れ領域であるため正確な推定が困難である。そこで、オクルージョン領域を推定するために物体の可視領域とオクルージョンが生じていない状態の物体領域を推定する IM R-CNN を提案する。図1(b)に示すように、OM R-CNN のオクルージョン領域推定部分をオクルージョンが発生していない状態の物体領域を推定するネットワークに置き換える。そして、推定結果を用いてオクルージョン領域を推定する。提案手法では、OM R-CNN で学習していたタスクに加え、物体の形状の推定を同時に学習する。

### 2.2 物体間の上下関係のツリー表現

多品種ばら積みのピッキングにおいて、オクルージョンが生じている物体は他物体が上に存在するため、物体の上下関係を把握して把持戦略を決定する必要がある。そのため、提案手法の推定結果をもとに物体間の上下関係をツリー構造で表現することでシーンに対する最適な把持順序を決定する。上下関係ツリーの例を図2に示す。上下関係ツリーの構築アルゴリズムを Algorithm1 に示す。各物体に対するオクルージョン領域の推定の結果を  $x_i$ 、画像中の全物体に対するオクルージョン領域の推定の結果を  $X = \{x_1, \dots, x_n\}$  とし、構築した上下関係ツリーを  $L$  とする。

## 3. 評価実験

OM R-CNN および IM R-CNN の有効性を調査するために、推定した可視領域およびオクルージョン領域を評価する。オクルージョン推定手法の学習に用いるデータセットとして、Amazon Robotics Challenge(ARC) Multi-task Dataset[2] を用いる。データセットのうち、880枚を訓練データ、220枚を検証データとする。また、推定結果をもとに構築した上下関係ツリーの評価を行う。

### 3.1 可視領域の推定

提案手法である OM R-CNN および IM R-CNN と従来手法の Mask R-CNN による可視領域推定精度を比較する。可視領域の推定における評価指標として、セグメンテーションタスクの一般的な評価指標である Global Accuracy(GA), Class Accuracy(CA), Mean IoU を用いる。可視領域の推定の評価結果を表1に示す。表1より、提案

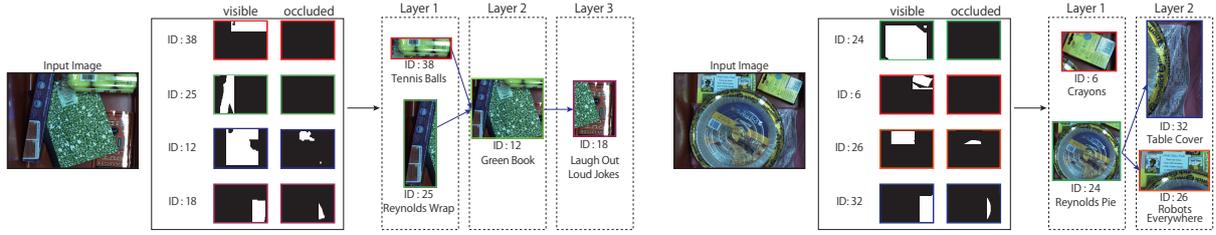


図 3：上下関係ツリーの構築結果例

表 1：可視領域およびオクルージョン領域推定の評価結果

	Object Area			Interpolation			Occluded Area			
	GA	CA	Mean IoU	GA	CA	Mean IoU	FN	FP	CA	Mean IoU
Mask R-CNN	76.7	61.8	54.0	-	-	-	-	-	-	-
OM R-CNN	<b>80.2</b>	65.3	<b>56.9</b>	-	-	-	<b>16.4</b>	<b>16.1</b>	<b>45.9</b>	<b>36.6</b>
IM R-CNN	79.4	<b>68.1</b>	56.1	<b>80.6</b>	<b>63.7</b>	<b>51.8</b>	27.5	25.7	28.0	17.9

#### Algorithm 1 Creating a vertical relationship layer

```

Input: list of occlusion segmentations of each object:  $X = \{x_1, \dots, x_n\}$ 
Initialize result layers:  $L \leftarrow \{\}$ 
Initialize a list for depth 1:  $l_1 \leftarrow \{\}$ 
for  $x_i \in X$  do
  if  $x_i$  does not contain occluded areas then
    Add  $x_i$  into  $l_1$ :  $l_1 \leftarrow l_1 \cup \{x_i, \text{None}\}$ 
    Remove  $x_i$  from  $X$ :  $X \leftarrow X \setminus \{x_i\}$ 
  end if
end for
Add  $l_1$  into  $L$ :  $L \leftarrow l_1$ 
Set depth  $d$  as 2:  $d \leftarrow 2$ 
while  $X \neq \emptyset$  do
  Initialize a list for depth  $d$ :  $l_d \leftarrow \{\}$ 
  for  $x_i \in X$  do
    if The occluded area is covered by objects in  $L$  then
      Find upper object  $x_j$  on  $x_i$ 
      Add  $x_i$  into  $l_d$ :  $l_d \leftarrow l_d \cup \{x_i, x_j\}$ 
      Remove  $x_i$  from  $X$ :  $X \leftarrow X \setminus \{x_i\}$ 
    end if
  end for
  Add  $l_d$  into  $L$ :  $L \leftarrow l_d$ 
   $d \leftarrow d + 1$ 
end while

```

手法は、すべての評価指標において Mask R-CNN を上回る精度であった。このことから、可視領域のセグメンテーションとオクルージョン領域のセグメンテーションには相関関係があり、同一ネットワーク内にオクルージョン領域推定を導入することで可視領域推定の精度を向上させることができる。

### 3.2 オクルージョン領域の推定

オクルージョン領域の推定における評価指標には CA と Mean IoU を用いる。また、オクルージョン領域推定タスクでは、小さな領域であっても物体のオクルージョン領域を正しく推定した場合、上下関係を正しく構築することが可能である。そのため、物体に対して推定したオクルージョン領域が正しいかを評価する False negative rate (FN) と False positive rate (FP) を用いる。FN を式 (1) に、FP を式 (2) に示す。

$$FN = \frac{\#of\ undetected\ objects}{\#of\ occlusion\ occur\ objects} \quad (1)$$

$$FP = \frac{\#of\ erroneously\ detected\ occlusion}{\#of\ total\ objects} \quad (2)$$

オクルージョン領域の推定の評価結果を表 1 に示す。表 1 より、OM R-CNN はすべての評価指標において、IM R-CNN を上回る精度であった。このことから、IM R-CNN は、オクルージョン領域推定の精度が可視領域推定および

オクルージョンの発生していない状態の物体領域推定の精度に依存してしまうため、低精度であったと考える。提案手法によるオクルージョン領域の推定において、CA および Mean IoU は低精度である。しかし、物体間の上下関係の構築において、小領域でもオクルージョン領域を正確に推定できれば上下関係を構築可能であるため、問題ないと考える。

### 3.3 上下関係ツリーの評価

提案手法の推定結果をもとに構築した上下関係ツリーの評価を行う。上下関係ツリーの評価には、構築した物体の上下関係の正答率を用いる。構築した物体の上下関係の正答率である Correct answer rate を式 (3) に示す。式 (3) における構築した上下関係が正しい物体数を  $T$  とする。

$$CR = \frac{T}{\#of\ total\ objects} \quad (3)$$

OM R-CNN の推定結果を用いて構築した上下関係ツリーは 68.8%、IM R-CNN の推定結果を用いて構築した上下関係ツリーは 41.9% の精度であった。OM R-CNN が可視領域及びオクルージョン領域を最も高い精度で推定可能であるため、最も高い精度で上下関係ツリーが構築できたと考えられる。OM R-CNN の推定結果をもとに構築した上下関係ツリーを図 3 に示す。図 3 より、OM R-CNN を用いて推定した可視領域とオクルージョン領域を用いることで全ての物体で正しい上下関係を構築できることが分かる。

### 4. おわりに

本研究では、物体検出、可視領域のセグメンテーション、オクルージョン領域のセグメンテーションのタスクを 1 つのモデルで実現する OM R-CNN および IM R-CNN、推定結果をもとに構築する物体間上下関係のツリー表現を提案した。今後の課題として、物体の可視領域およびオクルージョン領域推定の高精度化による上下関係ツリーの高精度化が考えられる。

#### 参考文献

- [1] K. He, *et al.*, “Mask R-CNN”, ICCV, 2017.
- [2] MPRG, “Amazon Robotics Challenge(ARC) Multi-task Dataset”, [Online] [http://mprg.jp/research/arc-dataset\\_2017\\_j](http://mprg.jp/research/arc-dataset_2017_j)

#### 研究業績

- [1] 稲垣雄介 等, “多品種ばら積みピッキングにおける物体間の上下関係の予測とデータセットの提案”, 日本ロボット学会学術講演会, 2019.
- [2] Y. Inagaki, *et al.*, “Detecting layered structures of partially occluded objects for bin picking”, IROS, 2019.