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

○稲垣雄介 荒木諒介 平川翼 山下隆義 藤吉弘亘 (中部大学)

1. はじめに

多品種ばら積みのピッキングにおいて，オクルージョンが発生した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．これを解決するには，物体間の上下関係を予測して把持戦略を決定する必要がある．そこで，本研究では，ばら積みされた物体間の上下関係を獲得するための新たなデータセットである ARC Multi-task Dataset を提案し，物体間の上下関係の予測を行う．

2. 関連データセット

多品種ばら積みのピッキングを対象としたデータベースを表 1 に示す．Grasping Dataset [3] は物体の把持可能な領域に対するセグメンテーションのデータ，Amazon Picking Challenge Object Scans(APCOS) [2] は RGB-D 点群に対するセグメンテーションデータが含まれている．JSK のデータセット [4] はインスタンスセグメンテーションラベルと各インスタンスに対する可視領域ラベルとオクルージョン発生領域ラベルがを含まれている．Grasping Dataset 及び APCOS は物体のオクルージョン領域にアノテーション情報がないため，物体間の上下関係の予測に用いることはできない．JSK のデータセットは画像合成で物体のばら積みシーンを生成しており，物体の見え方が限られていたり，本来あり得ない上下関係が存在する．本研究では，実環境における多品種ばら積みを対象とし，物体間の上下関係を評価を可能とするデータセットである ARC Multi-task Dataset を提案する．

3. ARC Multi-task Dataset

本研究では，図 1 に示すように物体領域及びオクルージョン発生領域を用いることで物体間の上下関係を獲得するためのデータセットを提案する．

3.1 データセット概要

提案する ARC Multi-task Dataset は，Amazon Robotics Challenge (ARC) 2017 の競技で用いられた 40 種類のアイテムから構成されており，各アイテムは箱状のアイテムやビニールで梱包されたアイテム，非剛体のアイテムなど様々な形状や属性を保持している．本データセットは，1280 × 960 pixels の RGB 画像，深度画像および 3D モデルを含む．1 シーンあたり 1 個から 8 個のアイテムが含まれており，様々なオクルージョンが発生するように乱雑に配置したシーンを含む．アイテムを複数個配置して撮影したシーンが 1,100 枚，また，全てのシーンに対して Bounding Box Label，Object Segmentation Label，Occluded-Area Segmentation Label の 3 つのラベルが付与されている．Bounding Box Label は，物体認識の学習及び評価に用いるための全てのシーンに対するバウンディングボックスの座標とアイテムの ID である．オクルージョンが発生している場合，見える範囲をバウンディングボックスで表す．Object Segmentation Label は，物体領域のセマンティックセグメンテーションラベルであり，ピクセル単位でアイテム ID に対応した色で表現されている．物体間の上下関係を予測するために，物体のオクルージョン発生領域をラベル付けした Occluded-Area Segmentation Label を導入する．Occluded-Area Segmentation Label は，

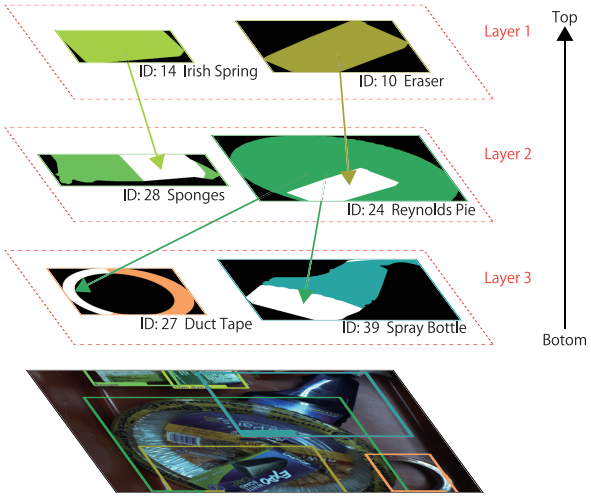


図 1 予測された物体間の上下関係ツリー

表 1 データセットの比較

Dataset	#Sample	Depth Image	#Class	#Bbox	#Semantic Segmentation Label	#Occluded Area Segmentation Label	#Grasping Label	Robotic Grasping Dataset	bin picking
Cornell University The Grasping Rectangle Dataset[1]	1035	✓	280	-	-	-	1035	✓	
UC Berkeley APCOS[2]	600		27	-	600	-	-	✓	✓
MIT Grasping Dataset(把持ハンド)[3]	778	✓	-	-	-	-	778	✓	✓
MIT Grasping Dataset(吸着ハンド)[3]	1837	✓	-	-	-	-	1837	✓	✓
JSK Laboratory Dataset[4]	∞		40	-	-	∞	-	✓	✓
ARC Multi-task Dataset	1500	✓	40	14,824	1500	14,824	-	✓	✓

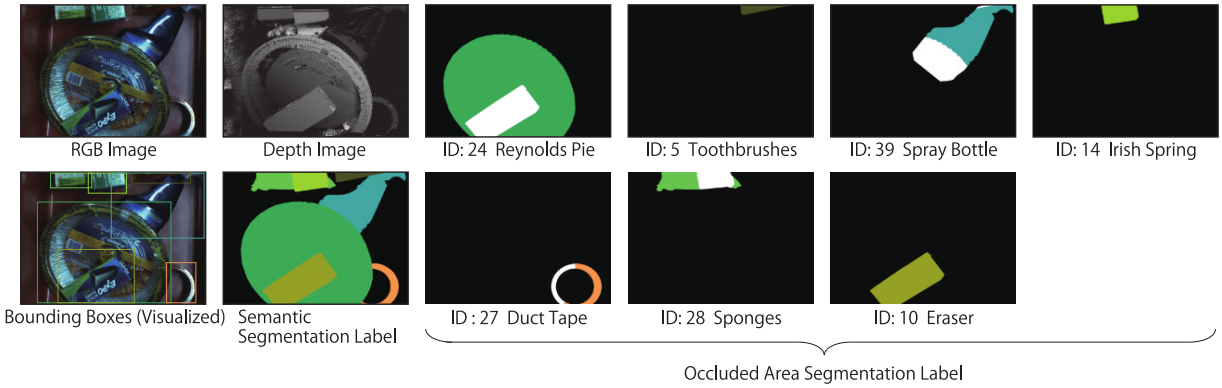


図 2　ARC Multi-task Dataset の画像サンプル

表 2　評価結果

	Mask			Occluded Area Mask	
	Global Accuracy	Class Accuracy	Mean IoU	Class Accuracy	Mean IoU
Mask R-CNN	76.7	61.8	54.0	-	-
Occluded-Area Mask R-CNN	80.2	65.3	56.9	45.9	36.6

図 2 のように各アイテムに対して物体領域とオクルージョン発生領域にラベルが付けられている．

3.2　物体間の上下関係のツリー表現

物体間の上下関係は，最も上にある物体を Layer1，その下にある物体を Layer2 と階層的にツリー表現で表す．これにより，物体に対してどの物体がオクルージョンしているかを認識することで把持順序の獲得が可能となる．物体の上下関係ツリーの例を図 1 に示す．このようなツリー表現を用いることで，物体間の上下関係の評価が可能となる．

4.　ベンチマーク評価

物体の上下関係ツリーを作成するには物体の位置，物体領域，オクルージョン発生領域を獲得する必要がある．物体に対するオクルージョン発生領域を獲得するために，インスタンスセグメンテーション手法である Mask R-CNN [5] に Occluded-Area Segmentation タスクを追加したネットワークを用いる．表 2 に，ベンチマークの結果を示す．Occluded-Area Mask R-CNN は Mask R-CNN と同等の精度で物体領域のセグメンテーションが可能である．また，各検出物体の Occluded-Area Segmentation では，Class Accuracy が 45.9%，Mean IoU が 36.6%であった．

Occluded-Area Mask R-CNN により獲得される Mask と Occluded-Area Mask を用いて作成した物体の上下関係ツリーの評価を行う．作成した上下関係ツリーの例を図 3 に示す．図 3 より，Occluded-Area Mask R-CNN を用いて作成した上下関係ツリーは，全ての物体で正しい上下関係を作成できている．Occluded-Area Mask R-CNN は，正しい上下関係を 68.8%の精度で獲得した．

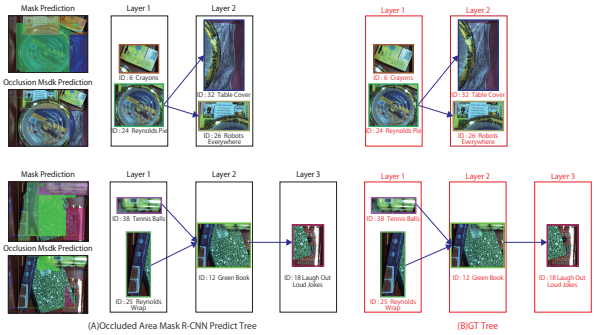


図 3　上下関係ツリーの出力結果例

5.　おわりに

本研究では，物体の上下関係を予測するための上下関係ツリーの作成手法の提案及び，物体間の上下関係を獲得するためのデータセット (ARC Multi-task Dataset) を提案した．評価の結果，上下関係ツリーを 68.8% の精度で作成することが可能であった．今後の展望として，Occluded-Area Segmentation 及び上下関係ツリーの作成の高精度化に取り組む．

謝辞

この成果は，国立研究開発法人新エネルギー・産業技術総合開発機構 (NEDO) の委託業務の結果得られたものである．

参　考　文　献

[1] I. Lenz, et al., “Deep Learning for Detecting Robotic Grasps,”RSS, 2013.
[2] UC Berkeley, “Amazon Picking Challenge Object Scans,” [Online] `http://rll.berkeley.edu/amazon\_picking\_challenge/`
[3] A. Zeng, et al., “Robotic Pick-and-Place of Novel Objects in Clutter with Multi-Affordance Grasping and Cross-Domain Image Matching,”ICRA, 2018.
[4] K. Wada, et al., “Instance Segmentation of Visible and Occluded Regions for Finding and Picking Target from a Pile of Objects,”JSAI, 2018
[5] K. He, et al., “Mask R-CNN,”ICCV, 2017.