The 33rd Annual Conference of the Japanese Society for Artificial Intelligence, 2019

物体インスタンスの重なりを考慮した双腕ロボットによる自律学習型ターゲットピッキングシステム

Self-supervised learning base target picking system for dual-arm robot considering object instance occlusion

北川 晋吾 Shingo Kitagawa　岡田 慧 Kei Okada　稲葉 雅幸 Masayuki Inaba

東京大学大学院 情報理工学系研究科
the University of Tokyo, Graduate School of Information Science and Technology

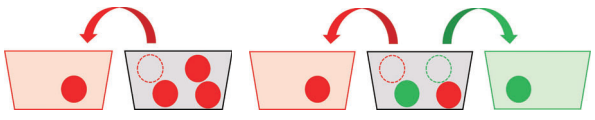
Recently, robots are introduced to warehouses and factories for automation and are expected to execute dual arm manipulation as human does. We focus on target picking task in the cluttered environment and aim to realize a robot picking system which the robot selects and executes proper grasping motion from single-arm and dual-arm motion. In this paper, we propose a self-supervised learning based target picking system with selective dual-arm grasping. In our system, a robot first learns how to grasp and how to distinguish items with synthesized dataset. The robot then executes and collects grasp trial experiences in the real world and retrains grasping model with the collected trial experiences. Finally, We also propose the learning based target picking system with selective dual-arm grasping and evaluate picking task experiments in the cluttered environment such as warehouse.

1. はじめに

近年，様々な分野でロボットの社会導入が推められている．特に中小工場や倉庫などの複雑環境でマニピュレーション作業を行うロボットシステムは，Amazon Robotics Challenge [Correll 16] など盛んに研究が行われており，特に学習手法を用いたシステムが高い作業成功率を示している [Pinto 16]．これらの研究では多様な物体が重なりあう複雑環境に対応するために，大量のデータセットを用いて環境認識モデルを学習する手法が採られているが，学習手法において動作の多様化に対するデータセット生成のコスト増加が問題となる．また既存の学習手法では単腕把持動作のみがピッキングタスクに導入されているのが現状である．本研究では上記の問題を解決しながら，複雑環境での学習型ロボットピッキングシステムに双腕マニピュレーション動作を導入することを目的とする．

2. 複雑環境におけるピッキングタスク

複雑環境におけるピッキングタスクは図 1 に示すランダムピッキングとターゲットピッキングに分類される．



(a) ランダムピッキング　(b) ターゲットピッキング

図 1: ランダムピッキングとターゲットピッキング

ランダムピッキングとは，図 1(a) に示すように環境に置かれた物品（ 図 1(a) 内の赤丸）を全て収納するタスクであり，これまで物品の意味ラベルに関係なく把持動作を行う研究が多く行われてきた [Pinto 16]．一方，ターゲットピッキングとは，図 1(b) に示すように環境内のターゲット物品（ 図 1(b) 内の赤丸）のみを収納するタスクであり，近年ランダムピッキングの学習手法の発展として研究が行われている [Wada 18]．このタスクはターゲット物品をいかに効率的に収納するかが課題となり，ターゲット物品を移動する妨げになる障害物物品（ 図 1(b) 内の緑丸）を他の場所に移動することができ，複雑環境で物品同士が重なりあった環境の中からターゲット物品を効率的に把持・収納する動作計画が必要になる．和田ら [Wada 18] は，複雑環境での物体の重なりあい（ 遮蔽領域）を認識しターゲット物品に重なっている物品（ 障害物物品）を移動することで，ターゲット物品を効率的に把持する動作計画を行う学習的認識手法を提案してされている．ランダムピッキングは，環境内の物品全てがターゲット物品というターゲットピッキングの一種であるため，本研究では拡張性の高いターゲットピッキングを対象タスクとして研究を行う．

連絡先: 北川 晋吾，東京大学大学院情報理工学系研究科，113-8656，東京都文京区本郷 7-3-1，s-kitagawa@jsk.imi.i.u-tokyo.ac.jp

3. 本研究の目的

本研究の目的は，複雑環境でのターゲットピッキングシステムに双腕マニピュレーション動作を導入することである．しかし，新たなマニピュレーション動作をピッキングシステムに導入する場合に主に2 つの問題が存在する．1 つ目の問題は，学習型ピッキングシステムに新たなマニピュレーション動作を新たに導入する場合，その動作の学習のためにより多くのデータセットが必要となる点である．またロボットのマニピュレーション動作に関しては，ロボット自身の実世界での動作試行なしで動作の評価は行えないため，ロボット実機による自律学習（ Self-supervised learning）が必要となる．

2 つ目の問題は，多様なマニピュレーション動作の中から状況に適した動作を評価・選択する必要がある点である．これは本研究で取り扱う単腕・双腕マニピュレーション動作それぞれの長所・短所を評価して，どちらが状況に適しているかをロボット自身が選択する必要がある．

したがって本研究では上記の2 つの問題点を解決した自律学習型双腕ターゲットピッキングシステムを実現することが目的となる．本研究で提案する双腕ターゲットピッキングシステムのシステム構成図は図 2 に示し，この提案システムを，自動データセット生成と把持点予測モデルの学習，把持点予測モデルの実世界把持経験への適応，選択的双腕把持を行うターゲットピッキングシステムの3 つに分けて説明する．

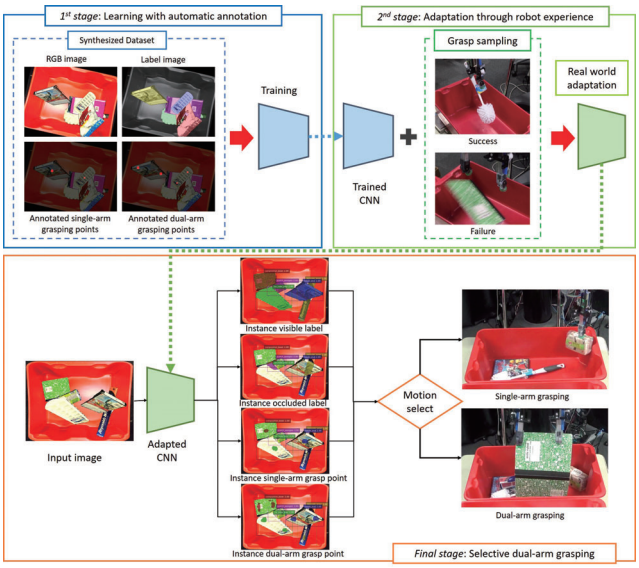
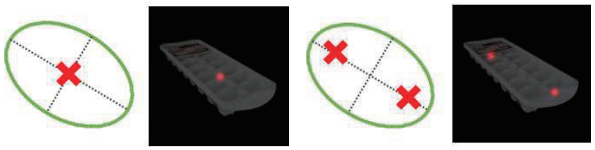


図 2: 提案する自律学習型双腕ターゲットピッキングシステム

4. 自動データセット生成と把持点予測モデルの学習

4.1 把持点予測データセットの自動生成

自動データセット生成のために本研究では物体形状に基づく把持点アノテーションアルゴリズムを設計する．本研究で提案する把持点アノテーションアルゴリズムでは，それぞれの対象物品に対して6 枚以下の RGB 画像を用いて，単腕・双腕把持における把持点を物体の幾何形状から図 3 に示すように計算する [Kitagawa 18].



(a) 単腕把持点の自動アノテーション (b) 双腕把持点の自動アノテーション

図 3: 単腕・双腕把持のための自動把持点アノテーション

把持点データセットの自動生成は複数の対象物品インスタンスの RGB 画像を変形させながら 1 枚の背景画像に貼り合わせることで，複数物品が重なりあい遮蔽しあう複雑環境の RGB 画像を自動生成する．また画像貼り合わせによって他物品によって覆いかぶさられた領域を物体インスタンスの遮蔽領域 occ としてアノテーションする．同様にして単腕・双腕把持の把持点に関してもアノテーションを行うが，他物品に対象物品領域の 10% 以上が覆いかぶさっている場合は物理的に対象物品が遮蔽されており把持不可能であるとみなす．

4.2 Mask-RCNN を元にした物体インスタンス可視・遮蔽領域分割を同時に行う把持点予測モデルの学習

本研究では以下に説明する把持点予測モデルと自動生成したデータセットを用いた学習手法を提案する．

この把持点予測モデルの全体構成は Mask-RCNN [He 17] をベースに構成されており，RGB 画像を入力として特徴量抽出器を介して得た特徴量を用いて，予測した複数の興味領域 r に対して物体領域矩形 b，意味ラベル c_{class}，ピクセル単位のインスタンス意味ラベル c_{inst}，把持可能ラベル c_{grasp} の予測を行う．インスタンス意味ラベル c_{inst} とは，物体領域矩形 b 内のあるピクセルが，物体インスタンスに属しているか vis，他物体との重なりで見えなくなっているか occ，もしくは背景であるか bg、を示すものである．

把持可能領域分割に関する損失 L_{single}, L_{dual} は把持戦略 s_{grasp} と興味領域 r に対して，r 内のピクセル (x, y) における c_{grasp} に関する予測出力 $h_{s_{grasp}}(r, x, y)$ が正解データ $l_{s_{grasp}}(r, x, y)$ を予測するように学習する．s_{grasp} と c_{grasp} についての r に関する重み $w^{c_{grasp}}_{s_{grasp}}(r)$ を式 (1) のように計算する．式 (1) において，$N_{foreground}(r)$ は正解ラベル画像における r 内の背景ラベルではないピクセル数，$N_{c_{grasp}}(r)$ は正解ラベル画像における r 内の把持ラベル c_{grasp} であるピクセル数，$\alpha_{c_{grasp}}$ は c_{grasp} についての定数である．

$$w^{c_{grasp}}_{s_{grasp}}(r) = \begin{cases} \frac{N_{foreground}(r)}{\alpha_{c_{grasp}} N_{c_{grasp}}(r)} & (N_{c_{grasp}}(r) \neq 0) \\ 0 & (N_{c_{grasp}}(r) = 0) \end{cases} \quad (1)$$

本研究では把持ラベルの定数としては自動データセットに対する学習では $\alpha_{graspable}$ は 1.0，$\alpha_{ungraspable}$ は 1.0 に設定する．そして s_{grasp} と r における把持可能ラベルに関する損失 $L_{s_{grasp}}(r)$ は、式 (1) を重みとした重み付きソフトマックス交差誤差 $WSCE$ で計算する．よって s_{grasp} における把持可能領域分割に関する損失 $L_{s_{grasp}}$ は前景フィルタ FG と $L_{s_{grasp}}(r)$ を用いて式 (2) に示すように計算する．

$$L_{s_{grasp}} = \frac{\sum_{r \in R} FG(l_{class}(r)) L_{s_{grasp}}(r)}{\sum_{r \in R} FG(l_{class}(r))} \quad (2)$$

把持点予測モデルの全体損失 L_{total} は興味領域予測の損失 L_{rpn}， 物体インスタンス検出の損失 L_{det}，物体インスタンス可視・遮蔽領域分割の損失 L_{inst}，把持可能領域分割の損失 L_{single}, L_{dual} の和として計算する．

5. 把持点予測モデルの実世界把持経験への適応

5.1 ロボット実機による実世界把持経験収集

前項で学習したモデルの出力を用いて実世界で把持経験を収集する．把持経験の収集工程の概要については図 4 に示すように，学習済み把持点予測モデルの出力を用いた重み付きランダムサンプリングにより把持点を決定し実行することで，ロボットが自動的に把持経験を収集する [Kitagawa 18].

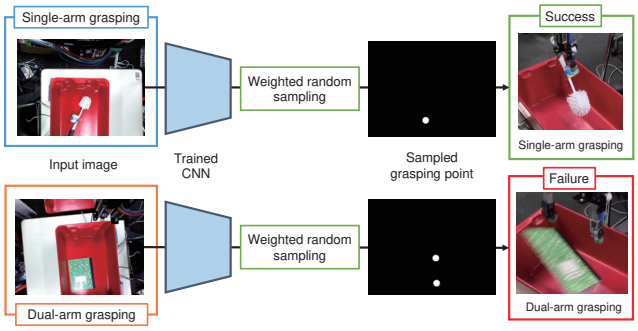


図 4: 把持点予測モデルを用いた実世界把持経験収集

5.2 再学習による実世界把持経験への適応

把持点予測モデルの予測結果における把持点の獲得と選定は，実世界での把持試行とそれによって収集されたデータセットを用いた把持点予測モデルの再学習によって実現される．再学習の際には，把持点予測モデルは意味ラベルに関しては学習済みであるとみなし，単腕・双腕把持の把持点に関する損失 L_{single}，L_{dual} の合計のみをネットワーク全体の再学習の損失 L_{total}^{adapt} として逆伝播する．

6. ターゲットピッキングのための選択的双腕把持動作システム

6.1 選択的双腕把持動作を行うターゲットピッキングシステムの動作フロー

本研究では選択的双腕把持動作を行いながらターゲットピッキングタスクを実行するための動作フローを設計する．まず 5. 節で再学習を行った把持点予測モデルを用いて，対象物品の画像から把持すべき物品と把持点，把持戦略を選択し，その選択した物品がターゲット物品であるかどうかをチェックする．そして，把持物品がターゲット物品である場合には，予測した把持点と把持戦略を用いて把持動作ののちに収納動作を実行し，ターゲット物品でない場合には，予測した把持点と把持戦略を用いて把持動作ののちに移動動作を実行する．

6.2 Mask-RCNN を元にした把持点予測モデルを用いた双腕把持動作選択システム

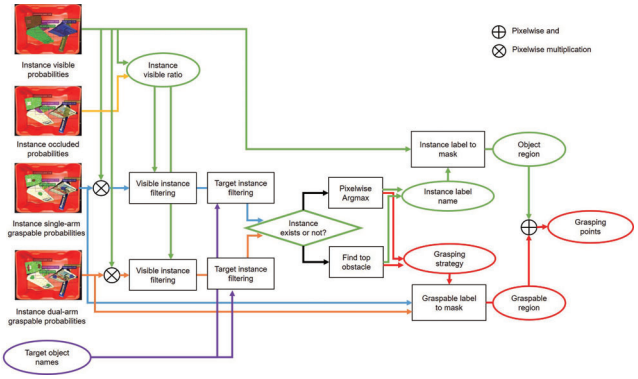


図 5: Mask-RCNN を元にした把持点予測モデルを用いた双腕把持動作選択システム

本研究では図 5 に示すような，Mask-RCNN を元にした把持点予測モデルを用いたの選択的双腕把持動作選択システムを提案する．まず各物体インスタンスの可視・遮蔽領域分割を行うことで，各物体インスタンスの可視度 $ratio_{vis}$ を計算する．ある物体インスタンス i についての可視度 $ratio_{vis}(i)$ は物体インスタンス i の可視領域内のピクセル数 $N_{vis}(i)$ を可視・遮蔽領域内のピクセル数の合計 $N_{vis}(i) + N_{occ}(i)$ で割ることで計算される．次に物体インスタンス i の可視度 $N_{vis}(i)$ について可視度の閾値 $threshold_{vis}$ を設定することで、この閾値以上の物体インスタンスを可視物体インスタンスとして選択する．選択した可視物体インスタンスに対して、意味ラベルの確率画像と把持可能ラベルの確率画像のピクセル積算を行うことで積算結果画像を生成し，この画像の最も高い値を示すピクセルを実行する把持戦略と実行する把持点として決定する．選択した可視物体インスタンスにターゲット物品が含まれていない場合には，遮蔽されているターゲット物品を 1 つ選択し、それを遮蔽している物品インスタンスについて、可視度が $threshold_{vis}$ を超える遮蔽物品インスタンスを探索していく。そして探索結果の遮蔽物品インスタンスを移動させる障害物インスタンスと設定し，ターゲット物品同様に，障害物物品に対する把持戦略と把持点を決定する．本研究では可視度の閾値 $threshold_{vis}$ を 0.9 と設定して動作選択を行う．

7. 検証実験

7.1 実験に用いる対象物品

本研究の検証実験で用いる対象9 物品は Amazon Robotics Challenge [Morrison 17] においてピッキングタスク対象物品として選ばれた 40 物品の中から選んだ．この9 物品はバインダー，ノートブック，袋入りくつした，製氷皿，アルミニウムホイル，DVD，スポンジ，テーブルクロス，トイレブラシである．この選んだ9 物品は，双腕把持しづらいが単腕把持では把持できる，単腕把持で把持しづらいが双腕把持では把持できる，どちらの単腕・双腕把持でも把持できる，の3 つのカテゴリに分類でき，各カテゴリには表 1 に示すようにそれぞれ2 から 4 物品が属する．

表 1: 対象9 物品の単腕・双腕把持に関する分類

Graspable category	Object names
Single-arm graspable	DVD, Toilet brush, Brown sponges
Dual-arm graspable	White binder, Green notebook
Both graspable	Ice cube tray, White socks, Aluminum foil, Pink table cloth

7.2 Mask-RCNN を元にした把持点予測モデルによる物体インスタンス可視・遮蔽領域分割結果

Mask-RCNN を元にした把持点予測モデルによる意味領域分割の評価として，クラス別平均平均適合率 mAP，クラス別平均領域分割精度 mSQ，クラス別平均インスタンス検出精度 mRQ，クラス別平均インスタンス領域分割精度 mPQ の4 指標 [Wada 18] を用いて評価を行った．人手でアノテーションした評価データセットに対して評価を行ったところ，表 2 上段に示す結果が得られ，自動生成データセットで学習したモデルでもクラス別平均領域分割精度 mSQ は 0.503 と高い値を示していることがわかる．また表 2 下段は人手でアノテーションして作成した比較用学習データセットを用いて学習したモデルの評価結果であり，自動生成データセットで学習したモデルと大差ない結果をしておりクラス別平均領域分割精度 mSQ も 0.004 だけ高い値を示すという結果が得られた．

表 2: Mask-RCNN を元にした把持点予測モデルによる物体インスタンス可視・遮蔽領域分割結果

Model	mAP	mSQ	mRQ	mPQ
Trained with synthesized dataset (Ours)	0.491	0.503	0.449	0.240
Trained with human annotated dataset	0.606	0.499	0.329	0.169

7.3 双腕ターゲットピッキングシステムの検証実験

本実験では再学習のために2 回の把持試行と再学習を行い，単腕把持 90 回，双腕把持 90 回の計 180 回の把持試行経験でモデルを再学習した．

7.3.1 対象1 物品に対する選択的双腕把持の検証

本実験では前項と同様の選択的双腕把持の実験を行った．実験結果については表 3 に示す．表 3 に示すように再学習前と後では把持成功率が 52.2% と 84.4% と大きく向上した．単腕・双腕把持の成功率についても同様に向上しており，以上の結果から把持経験収集による再学習を行うことで把持成功確率が向上することを確認した．

表 3: Mask-RCNN を元にしたモデルによる把持結果

	Single-arm success	Dual-arm success	Total success
Before retraining	32 (68.1%)	15 (34.9%)	47 (52.2%)
After retraining	**65 (91.5%)**	**11 (57.9%)**	**76 (84.4%)**

7.3.2 対象3 物品に対するターゲットピッキングの検証

本項では対象3 物品に対するターゲットピッキングの検証実験について説明する．各物品1 つずつについてターゲットピッキングを実行する際に，他8 物品の中から2 物品を障害物としてターゲット物品の上に重ねることで複雑環境を作成した．この複雑環境を計9 シーン作成して実験を行い，各シーン2 回のタスク実行，計 18 回のターゲットピッキングを行った．ターゲットピッキング結果は表 4 に示すように，本研究で提案するターゲットピッキングシステムは 18 回中 11 回成功の 61.1% と高いタスク成功率を示している．

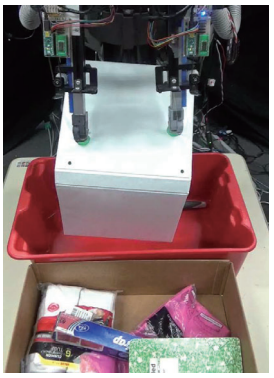
表 4: 提案手法によるターゲットピッキング結果

Success	Item drop	Mis-recognition	Mis-grasp	Obstacle removal failure	Total
11 (61.1%)	1 (5.6%)	2 (11.1%)	1 (5.6%)	2 (11.1%)	18

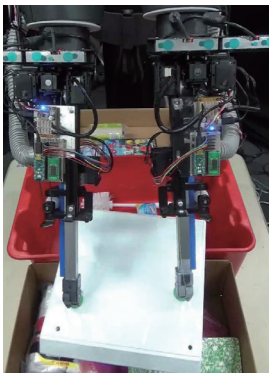
7.3.3 複雑環境における Mask-RCNN を元にした把持点予測モデルを用いたターゲットピッキング

Mask-RCNN を元にした把持点予測モデルによる最終実験として，再学習した把持点予測モデルとその出力に基づく選択的双腕把持動作を用いて，対象9 物品が重なりあう複雑環境における選択的双腕把持動作を行うターゲットピッキングを行った．このときバインダー，ノートブック，袋入りくつした，アルミホイル，トイレブラシの5 物品をターゲット物品として，その他の物品を障害物物品として設定し，全ての物品を籠のなかにランダムに重ねて配置してターゲットピッキングを行った．またこの実験では物品の認識や把持が困難でタスク実行が終了できない場合が発生したため，その際には人が籠の中にある物品を重ねなおしてタスクを再開した．

本実験では2 シーンの複雑環境でターゲットピッキングを行った．両シーンともにテーブルクロスをターゲット物品と誤認識して収納しているが，それ以外は全てのターゲット物品を正しく収納している．また両シーンでともにバインダーとノートブックの把持の際に双腕把持動作を選択して行い，その他の物品に対しては単腕把持動作を行った．また図 6 に示すように，複雑環境においてロボットはターゲット物品であるバインダーを正しくターゲット物品であると認識し，かつ安定的に把持できる双腕把持動作とその把持点を適切に選択することで，把持・収納動作を問題なく実行した．さらに図 7 では，ロボットは DVD をターゲット物品を把持する際の障害物であると認識し，かつ安定的に把持できる単腕把持動作とその把持点を適切に選択することで，把持・移動動作を問題なく実行した．



(a) バインダー(ターゲット物品）の双腕把持動作



(b) バインダー(ターゲット物品）の双腕収納動作

図 6: 複雑環境でのターゲットピッキングにおけるターゲット物品の双腕把持・収納動作

(a) DVD(障害物物品）の単腕把持動作

(b) DVD(障害物物品）の単腕移動動作

図 7: 複雑環境でのターゲットピッキングにおける障害物物品の単腕把持・移動動作

8. 結論

本研究では，ロボットが状況に応じて選択的双腕把持を行う自律学習型ピッキングタスクシステムを実現した．今後は収集した実世界把持経験の偏りへの対応や多様な双腕マニピュレーション動作への拡張を行っていく．

参考文献

[Correll 16] Correll, N., et al.: Lessons from the Amazon Picking Challenge, *CoRR*, Vol. abs/1601.05484, (2016)

[He 17] He, K., et al.: Mask R-CNN, in *ICCV*, pp. 2980–2988 (2017)

[Kitagawa 18] Kitagawa, S., et al.: Multi-stage Learning of Selective Dual-arm Grasping Based on Obtaining and Pruning Grasping Points Through the Robot Experience in the Real World, in *IROS*, pp. 7123–7130 (2018)

[Morrison 17] Morrison, D., et al.: Cartman: The low-cost Cartesian Manipulator that won the Amazon Robotics Challenge, *CoRR*, Vol. abs/1709.06283, (2017)

[Pinto 16] Pinto, L. and Gupta, A.: Supersizing self-supervision: Learning to grasp from 50K tries and 700 robot hours, in *ICRA*, pp. 3406–3413 (2016)

[Wada 18] Wada, K., et al.: Instance Segmentation of Visible and Occluded Regions for Finding and Picking Target from a Pile of Objects, in *IROS*, pp. 2048–2055 (2018)