

アノテーションデータを活用したデータ拡張によるインスタンスセグメンテーションを用いた組立進捗推定

湯本茂樹[†], 三浦一真[†], Sarthak Pathak[‡],

Alessandro Moro[‡], 梅田和昇[‡]

○ Shigeki YUMOTO[†], MiuraKAZUMA[†], Sarthak PATHAK[‡],

Alessandro MORO[‡], Kazunori UMEDA[‡]

[†]: 中央大学大学院理工学研究科精密工学専攻, yumoto@sensor.mech.chuo-u.ac.jp

[‡]: 中央大学理工学部精密機械工学科

[§]: RITECS Inc.

＜要約＞ 本研究では、インスタンスセグメンテーションを用いて工場での製品組立進捗を推定する手法を提案する。インスタンスセグメンテーションにより部品を検出し、検出した部品の数や種類により進捗推定を行う。実利用においては多種多様な学習データの用意は困難であるため、物体に対して背景が固定されてしまい検出精度に影響を及ぼす。そこで、アノテーションデータを活用したデータ拡張により部品が完全に隠れないように遮蔽を制御することで、部品の検出精度向上を目指す。工場製品を模擬したカラーボックスに対して実験を行い、有効性を確認した。

＜キーワード＞ 工場, 深層学習, インスタンスセグメンテーション, データ拡張

1. 序論

近年, AI, IoT 技術や画像処理技術の発展によりスマートファクトリー化が進められている[1]. 中でも製品の組立進捗管理の自動化は業務の効率化や作業の問題発見などにおいて重要な課題である. しかし多品種少量生産型の工場では製品の組立が人による作業が中心となる. 人手による進捗管理は手間がかかるため作業進捗の共有化が進んでいない. そのため作業の遅れなどを把握できず納期に間に合わないことがあるといった問題がある. そこで工場内に設置された定点カメラから取得した映像から自動で組立進捗を推定する手法が有効であると考えられる.

組立進捗の推定手法の1つに作業者の手の動きを利用したものがある[2]. しかし, 組立対象が大きい場合は作業者によって組立動作が異なるため利用が難しい. またクラス分類による研究[3][4]が行われている. これらは作業工程を予めいくつかのクラスとして設定し, カメラの取得画像からクラス分類を行う. しかし近いクラスは変化が小さく見た目が似ているため誤判定が起きる場合があるという課題があ

る. また進捗推定にインスタンスセグメンテーションを用いた手法[5]が提案されている. インスタンスセグメンテーションとは物体検出と領域分割の両方を行うタスクであり, インスタンスセグメンテーションにより検出された部品の種類や数によって進捗を推定する手法である. しかし実利用においては十分な学習データを集めることが難しく, 背景の固定やオクルージョンによりセグメンテーションの検出精度が低下するという問題がある. 少ない学習データで精度を向上させる手法としてデータ拡張がある. 既存のデータ拡張手法として学習データの一部を遮蔽する Cutout[6], Random Erasing[7], Gridmask[8]が提案されている. しかし製品組立においては学習時に部品が完全に隠れてしまい検出精度が低下してしまう問題がある.

本研究では, アノテーションデータを活用したデータ拡張による組立進捗推定手法を提案する. 学習データを比較的小さなマスクでランダムに複数隠し, 学習データを増やすことで検出精度を向上させる. この際にアノテーションデータを活用することで矩形領域によって部品が遮蔽される割合を制御する.

2. 提案手法

2.1 提案手法概要

提案手法の流れを図1に示す。まず組立工程における製品状態の定義を行う。組立工程を任意の数に分割し、それぞれの組立工程における製品状態を定義する。具体的にはその組立工程における取り付けられている部品の種類や数を設定する。次にモデルに部品画像の学習を行う。その際にデータ拡張を適用する。認識時に学習したモデルに画像を入力し部品の検出結果を出力する。検出結果から進捗の推定を行う。

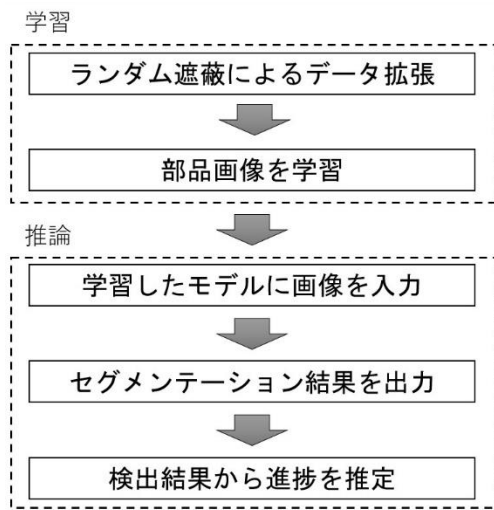


図1 提案手法概要

2.2 インスタンスセグメンテーション

インスタンスセグメンテーションとは物体検出と領域分割の両方を行うタスクである。インスタンスセグメンテーションには YOLACT[9]を使用する。YOLACT とは高速なインスタンスセグメンテーションが可能なモデルであり、物体検出と識別を同時に行う One-stage 型であるため、Two-stage 型の Mask-RCNN[10]と異なりリアルタイムでの検出が可能である。

2.3 進捗推定手法

組立進捗の推定はインスタンスセグメンテーションの検出結果を利用して行う。インスタンスセグメンテーションには YOLACT[9]を利用する。検出された部品の種類や数の組み合わせから条件分岐によって推定された進捗を出力する。図2に例を示す。検出結果が組立工程に当てはまらない場合は、誤った組立としてエラーを出力する。

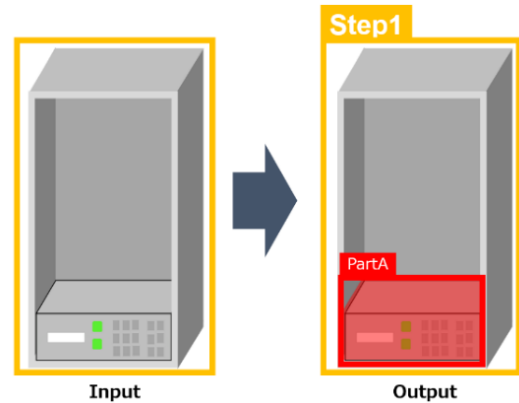


図2 進捗推定の例

2.4 アノテーションデータを活用したデータ拡張

一般的にオクルージョンに強くするデータ拡張として画像を矩形領域などで遮蔽する手法がある。Random Erasing[7]はランダムな大きさの矩形領域で遮蔽を行う。Gridmask[8]は規則的に並べられた正方形の領域でマスクを行う。しかし遮蔽によって検出対象が隠れてしまい検出精度が低下してしまう問題がある。そこでアノテーションデータを活用したデータ拡張を提案する。

遮蔽のアルゴリズムは Random Erasing[7]を利用する。まず矩形領域面積 S_e を s_l から s_h の範囲で設定する。次に矩形領域のアスペクト比 r_e を r_1 から r_2 の範囲で設定する。 S_e と r_e から矩形領域の縦幅 H_e と横幅 W_e を求める。

$$H_e = \sqrt{S_e \times r_e} \quad (1)$$

$$W_e = \sqrt{\frac{S_e}{r_e}} \quad (2)$$

最後に矩形領域を重畳する位置 x_e , y_e をランダムに設定する。矩形領域 I_e は (x_e, y_e) から $(x_e + W_e, y_e + H_e)$ で重畳する。矩形領域の各ピクセル値は0から255の間でランダムに割り当てられる。重畳する位置がアノテーションと重なっている場合はアノテーションされているセグメンテーションマスクとの重複面積 S_o を求める。セグメンテーションマスク面積 S_s との重複割合があらかじめ設定した値以下ならそのまま重畳し、超えていた場合は別の矩形領域を作成する。セグメンテーションマスクとの重複割合を計算することで矩形領域によって部品が完全に遮蔽されてしまうことを防ぐ。上記の処理を設定したマスクの数 n だけ繰り返す。マスクの例を図3に示す。



図3 アノテーションデータによるデータ拡張

3. 実験

3.1 実験環境

提案手法を用いて組立工程の推定実験を行った。組立対象としては現場での組立対象を模擬した2つのカラーボックスを積み重ねたものを使用した。作業動画はWebカメラによって取得した。使用したカメラはアイ・オー・データ機器社のネットワークカメラ Qwatch の TSWR-LP で、画素値は1980×1080に設定した。実験環境を図4に示す。



図4 実験環境

3.2 実験条件

学習データは332枚、テストデータは220枚、学習回数は200000iteration、バッチサイズは4とした。推定する組立工程を図5に示す。組立工程は0から9の10工程とした。検出する部品は、仕切り板、バケット、本、アルコールジェル、タンブラー、布とした。

提案手法のパラメータは矩形領域の数100、データ拡張の適用確率0.3、矩形領域のアスペクト比0.5~2.0、矩形領域面積0.002~0.01、セグメンテーションマスクとの重複割合0.1に設定した。これらの値は実験的に定めた。テストデータに対して正しく進捗推定ができるかを評価する。YOLACTのみ、YOLACTにRandom Erasing, Gridmask, 提案手法を適用したものの計4つを比較した。それぞれ5回学習をし、精度の平均値を算出した。

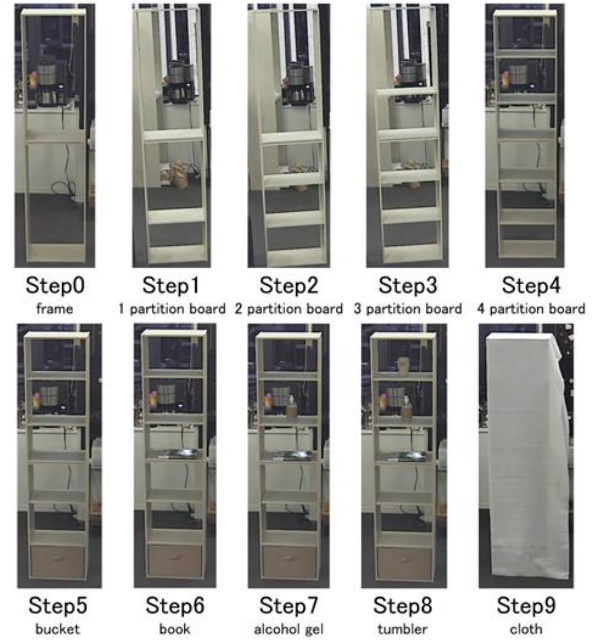


図5 進捗の設定

3.3 実験結果

実験結果を表1に示す。またそれぞれの混同行列を図6、図7、図8、図9に示す。提案手法が最も高い精度となった。既存のデータ拡張手法であるRandomErasingとGridmask、提案手法を適用したモデルはYOLACT単体より高い精度となった。このことからデータ拡張により汎化性能が高まり検出精度が上がったと考えられる。また既存のデータ拡張より高い精度が得られた。アノテーション箇所への遮蔽割合を制御することで、部品が完全に隠れてしまうことが減少したため精度が向上したと考えられる。YOLACT単体と提案手法の混同行列を比較すると、Step6の精度が14%向上している。既存のデータ拡張手法と比較しても提案手法が最も高い精度となっている。Step6は本を設置する工程である。本は他の部品と比較して画角によって見え方が変化するやすい。そのため複数の遮蔽によって様々なオクルージョンを再現することで既存のデータ拡張手法よりも高い精度が得られたと考えられる。

表1 実験結果

Data augmentation	Accuracy(%)
Yolact	80.5
Yolact+RandomErasing	82.1
Yolact+Gridmask	81.0
Yolact+Proposed method	83.0

0	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
1	0.03	0.92	0.02	0.03	0.00	0.00	0.00	0.00	0.00	0.00	0.00
2	0.00	0.09	0.91	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
3	0.00	0.08	0.25	0.64	0.03	0.00	0.00	0.00	0.00	0.00	0.00
4	0.00	0.00	0.02	0.19	0.77	0.00	0.00	0.00	0.00	0.00	0.02
5	0.00	0.00	0.01	0.04	0.09	0.70	0.00	0.00	0.00	0.00	0.16
6	0.00	0.00	0.00	0.00	0.00	0.08	0.64	0.00	0.00	0.00	0.28
7	0.00	0.00	0.00	0.00	0.00	0.00	0.67	0.00	0.00	0.00	0.33
8	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.65	0.00	0.00	0.35
9	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00	0.00	0.00
E	0.01	0.01	0.01	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.95
	0	1	2	3	4	5	6	7	8	9	E

図 6 混同行列 (YOLACT)

0	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
1	0.02	0.93	0.01	0.04	0.00	0.00	0.00	0.00	0.00	0.00	0.00
2	0.00	0.09	0.91	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
3	0.00	0.06	0.24	0.67	0.03	0.00	0.00	0.00	0.00	0.00	0.00
4	0.00	0.00	0.01	0.17	0.79	0.00	0.00	0.00	0.00	0.00	0.03
5	0.00	0.00	0.01	0.03	0.13	0.68	0.00	0.00	0.00	0.00	0.15
6	0.00	0.00	0.00	0.00	0.00	0.02	0.78	0.00	0.00	0.00	0.20
7	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.70	0.00	0.00	0.30
8	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.73	0.00	0.27
9	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00	0.00
E	0.00	0.00	0.04	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.95
	0	1	2	3	4	5	6	7	8	9	E

図 9 混同行列 (提案手法)

0	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
1	0.01	0.93	0.02	0.04	0.00	0.00	0.00	0.00	0.00	0.00	0.00
2	0.01	0.10	0.88	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00
3	0.00	0.08	0.14	0.76	0.02	0.00	0.00	0.00	0.00	0.00	0.00
4	0.00	0.01	0.04	0.17	0.74	0.00	0.00	0.00	0.00	0.00	0.04
5	0.00	0.00	0.01	0.02	0.15	0.71	0.00	0.00	0.00	0.00	0.11
6	0.00	0.00	0.00	0.00	0.00	0.02	0.72	0.00	0.00	0.00	0.26
7	0.00	0.00	0.00	0.00	0.00	0.00	0.65	0.00	0.00	0.00	0.35
8	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.72	0.00	0.00	0.28
9	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00	0.00	0.00
E	0.01	0.00	0.03	0.01	0.02	0.01	0.00	0.00	0.00	0.00	0.92
	0	1	2	3	4	5	6	7	8	9	E

図 7 混同行列 (Random Erasing)

0	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
1	0.04	0.91	0.01	0.04	0.00	0.00	0.00	0.00	0.00	0.00	0.00
2	0.00	0.13	0.87	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
3	0.00	0.06	0.23	0.69	0.02	0.00	0.00	0.00	0.00	0.00	0.00
4	0.00	0.03	0.05	0.20	0.66	0.00	0.00	0.00	0.00	0.00	0.06
5	0.00	0.00	0.01	0.08	0.04	0.68	0.00	0.00	0.00	0.00	0.19
6	0.00	0.00	0.00	0.00	0.00	0.02	0.73	0.00	0.00	0.00	0.25
7	0.00	0.00	0.00	0.00	0.00	0.00	0.70	0.00	0.00	0.00	0.30
8	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.74	0.00	0.00	0.26
9	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00	0.00	0.00
E	0.00	0.01	0.05	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.93
	0	1	2	3	4	5	6	7	8	9	E

図 8 混同行列 (Gridmask)

4. 結論

本研究では、アノテーションデータを活用したデータ拡張によるインスタンスセグメンテーションによる組立進捗推定手法を提案した。工場製品を模擬したカラーボックスに対して実験を行い、従来のデータ拡張手法を上回る精度を確認した。

今後の展望としてセグメンテーションマスクとの重複割合の変化や Salient Object Detection を活用した手法が考えられる。

参考文献

- [1] Wang, J., Ma, Y., Zhang, L., Gao, R. X., & Wu, D, "Deep learning for smart manufacturing", Methods and applications. Journal of manufacturing systems, 48, 144-156, 2018
- [2] M. Funk, A. Bachler, L. Bachler, T. Kosch, T. Heidenreich, "Working with augmented reality? A long-term analysis of in-situ instructions at the assembly workplace." Proceedings of the 10th international conference on pervasive technologies related to assistive environments. 2017. p. 222-229.
- [3] 大島 宏友, 白川 悠太, 吉井 崇哲, 蚊戸 健浩, 前川 卓也, 浪岡 保男: "深層学習を用いた製品状態の推定に基づく作業進捗の自動抽出." 生産システム部門講演会講演論文集 2022. 一般社団法人 日本機械学会, 2022. p. 206.
- [4] 橘川 拓実, 湯本 茂樹, Sarthak Pathak, Alessandro Mor, 梅田 和昇: "オクルージョンを考慮した深層距離学習による組立作業の進捗段階の推定", 動的画像処理実利用化ワークショ

- ップ DIA2023, IS1-15, pp.107-113, 2023.3.
- [5] 湯本 茂樹, 橘川 拓実, Sarthak Pathak, Alessandro Moro, 梅田 和昇: “インスタンスセグメンテーションを用いた組立進捗推定”, 動的画像処理実利用化ワークショップ DIA2023, IS1-13, pp.99-102, 2023.3.
 - [6] DeVries, Terrance, and Graham W. Taylor. "Improved regularization of convolutional neural networks with cutout." arXiv preprint arXiv:1708.04552 (2017).
 - [7] Zhong, Z., Zheng, L., Kang, G., Li, S., & Yang, Y. "Random erasing data augmentation." Proceedings of the AAAI conference on artificial intelligence. Vol. 34. No. 07. 2020. p. 13001-13008.
 - [8] Chen, P., Liu, S., Zhao, H., & Jia, J. "Gridmask data augmentation." arXiv preprint arXiv:2001.04086 (2020).
 - [9] Bolya, D., Zhou, C., Xiao, F., & Lee, Y. J. "YOLACT: Real-time instance segmentation." Proceedings of the IEEE/CVF international conference on computer vision. 2019. p. 9157-9166.
 - [10] He, K., Gkioxari, G., Dollár, P., & Girshick, R. "Mask r-cnn." Proceedings of the IEEE international conference on computer vision. 2017. p. 2961-2969.