

背景情報を考慮したインスタンスセグメンテーションによる

組立進捗推定

○湯本 茂樹 (中央大学), 三浦 一真 (中央大学), Alessandro Moro (RITEC Inc.),
Sarthak Pathak (中央大学), 梅田 和昇 (中央大学)

Assembly progress estimation by instance segmentation considering background information

○Shigeki YUMOTO (CHUO UNIVERSITY), Kazuma MIURA (CHUO UNIVERSITY),
Alessandro MORO (RITEC Inc.), Sarthak PATHAK (CHUO UNIVERSITY),
and Kazunori UMEDA (CHUO UNIVERSITY)

Abstract: This paper proposes a method for estimating product assembly progress in a factory using instance segmentation. Parts are detected by instance segmentation and progress is estimated based on the number and type of parts detected. The problem is that with small amounts of training data, background information and occlusion reduce the accuracy of segmentation detection. Data augmentation that multiple small random erasing confirmed the improvement in estimation accuracy.

1. 緒言

近年、画像処理技術や AI, IoT 技術の発展により工場の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という問題がある. 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組立動作が決まっていないため利用が難しい.

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てしまうという問題がある. またインスタンスセグメンテーションを用いた進捗推定手法[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る. 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情報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マスクする Cutout[6]や 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, ランダムマスクによるデータ拡張を利用した組立進捗推定手法を提案する. そこで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.

2. 提案手法

2.1 提案手法概要

提案手法の流れを Fig. 1 に示す. インスタンスセグメンテーションモデルに部品画像の学習を行う. その際にランダムマスクによるデータ拡張を適用する. 次に学習したモデルに画像を入力し検出結果を出力する. 検出結果から進捗の推定を行う.

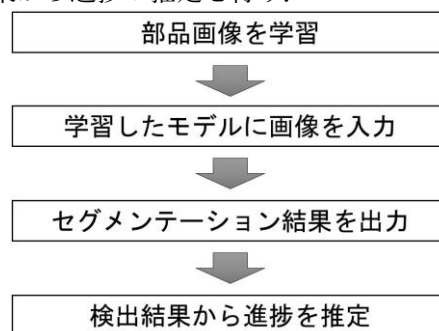


Fig. 1 Flow of the proposed method

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る. 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。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。Fig. 2に例を示す。

検出結果が進捗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。

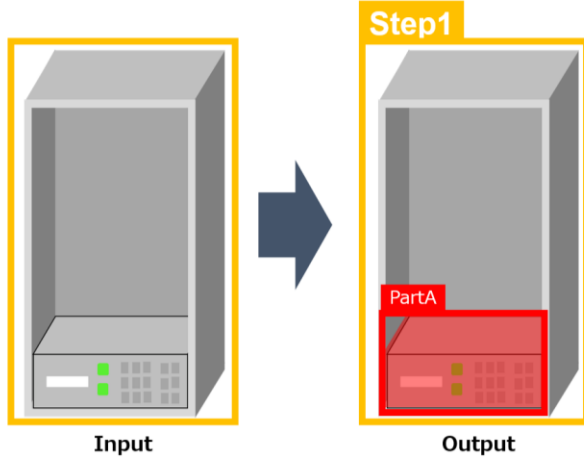


Fig. 2 Examples of progress estimation

2.4 複数のランダムマスクによるデータ拡張

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に矩形領域などでマスクをする手法が挙げられる。Random Erasing[7]はランダムな大きさの矩形領域でマスクを行う。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。しかしマスク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。そこで複数の小さなマスクでのデータ拡張を提案する。



Fig. 3 Example of random mask

マスクの例を Fig. 3 に示す。マスクのアルゴリズムは Random Erasing[7]を利用する。まずマスク面積 S_e は s_l から s_h の範囲で設定される。次にマスクのアスペクト比 r_e を r_1 から r_2 の範囲で設定される。 S_e と r_e からマスクの縦幅 H_e と横幅 W_e を求める。

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

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

最後にマスクを重畳する位置 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。上記のアルゴリズムを設定したマスクの数 n だけ繰り返す。

3. 実験

3.1 実験環境

提案手法を用いて組立進捗の推定実験を行った。組立対象としては半導体製造装置を模擬した2つのカラーボックスを積み重ねたものを使用した。Webカメラによって作業動画を取得する。作業動画の取得に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 を利用し、画素値は1980×1080に設定した。実験環境を Fig. 4 に示す。

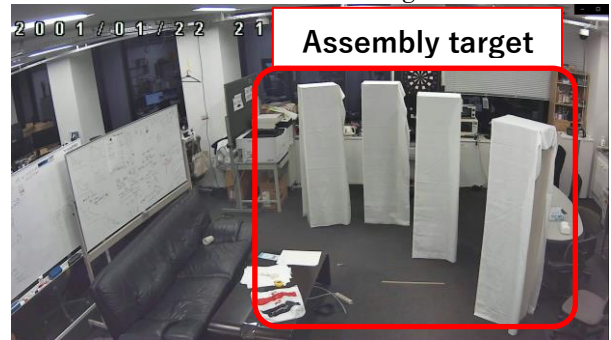


Fig. 4 Experimental environment

3.2 実験条件

学習データは332枚、テストデータは200枚、学習回数は200000iteration、バッチサイズは4とした。推定する進捗を Fig. 5 に示す。進捗は0から9の10工程とする。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。

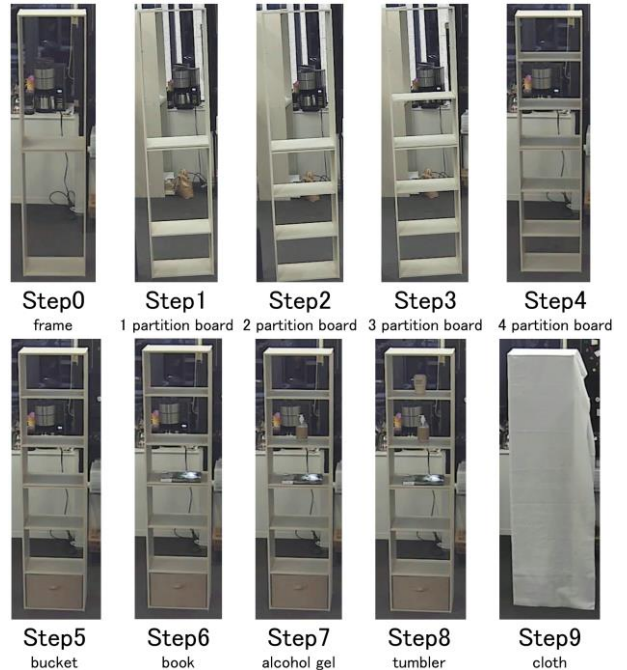


Fig. 5 Step of color box assembly progress

提案手法のパラメータはマスクの数 100, マスク確率 0.3, マスクのアスペクト比 0.5~2.0, マスク面積 0.002~0.01 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[7], Gridmask[8], 提案手法を適用したものと比較を行った。

3.3 実験結果

実験結果を Table 1 に示す。提案手法と Random Erasing を比較するとどちらも Yolact 単体より精度が向上している。精度が向上した理由としてデータ拡張により汎化性能が高まったためと考えられる。Random Erasing と比較して提案手法の方が精度向上している。このことからカラーボックスでは小さなマスクがより適していると考えられる。また Gridmask は Yolact 単体より精度が低下した。精度が下がってしまった原因として Gridmask は規則的に正方形でマスクを行うため、汎化性能が低下してしまったと考えられる。

Yolact+提案手法の混同行列を Fig. 6, Yolact のみを Fig. 7 に示す。混同行列を比較すると Step6 の精度が向上している。Step6 は本を設置する工程である。向上した理由として本が最も背景情報の影響を受けやすいためだと考えられる。本は平置きなのでカメラからの視点だと細長く見えるためアノテーションの際に背景を巻き込みやすい。また同じく細長い仕切り板と違い本は常に同じ場所に置かれるため背景が固定されやすい。これらが提案手法によって改善されたと思われる。

Table 1 Comparison of accuracy with existing methods

Data augmentation	Accuracy(%)
Yolact	83.0
Yolact+RandomErasing[7]	83.5
Yolact+Gridmask[8]	82.0
Yolact+Proposed method	85.0

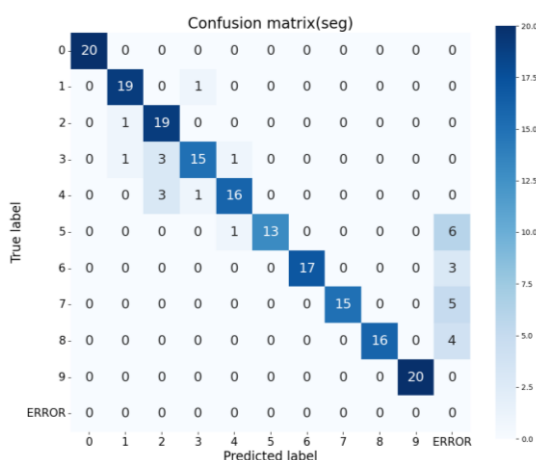


Fig. 6 Yolact+Proposed method

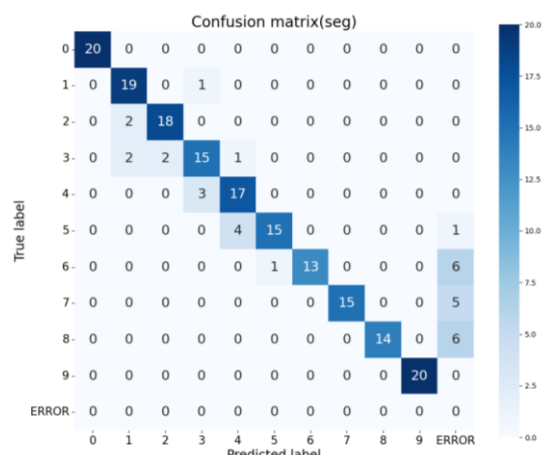


Fig. 7 Yolact only

4. 結言

本研究では、データ拡張を活用した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。

今後の展望としてランダムではなく背景を優先的にマスクする手法の改良が考えられる。

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.