

局所的な特徴と大域的な特徴に基づく組立進捗推定手法の融合 に向けた比較検討

○正津 隼光（中央大） Sarthak Pathak（芝浦工業大） 梅田 和昇（中央大）

本研究では、組立進捗推定に対し、局所的な特徴を重視した部品検出手法と製品全体の外観から大域的な特徴を抽出する深層距離学習手法とのハイブリッド化に向けた基礎検討として、両手法を比較検討する。両手法における学習の安定性や連続フレームにおける推定精度や変動特性を時系列的に評価し、両者の補完関係を明らかにすることで、局所と大域の特徴を融合した高精度かつ頑健な組立進捗推定手法の実現を目指す。

1. 序論

近年、AI・IoT技術の進展とともに工場のスマートファクトリー化が進展し、製品の組立作業における進捗の自動把握と可視化は、生産性の向上や作業遅延の早期発見の観点から極めて重要な課題となっている [1]。特に多品種少量生産が行われる工場現場では、人手による作業が多くなるため、進捗管理の自動化が強く求められている。

この課題に対して、近年では定点カメラから取得される映像を用いて製品の進捗を自動推定する手法が注目されている。その中でも特に、対象となる製品の視覚的特徴のどの側面に着目するかによって、大きく2つのアプローチに分類される。1つ目は、製品を構成する部品単位（局所の特徴）に着目した手法であり、インスタンスセグメンテーションなどを用いて部品単位の検出を行い、検出した部品の数や種類から進捗を推定する方法である [2]。このアプローチは、部品単位での構成情報を活用することで、工程ごとの変化を明確に捉えられる利点がある。一方で、背景の固定化やオクルージョンの影響により検出精度が低下することや、十分な学習データを確保することが困難であるといった課題も存在する。2つ目は、製品全体の外観（大域の特徴）に着目した手法であり、深層距離学習を用いて各進捗状態を特徴空間上で表現・分類する方法である [3]。この手法は、部品単位の検出に依存せず、製品全体の状態変化を柔軟に捉えられるため、オクルージョンやノイズに対して頑健性を持つ。一方で、工程間での外観変化が小さい場合には境界の判定が曖昧になり、誤判定が多くなるという課題もある。

このように局所の特徴と大域の特徴という観点からは、各々異なる強みと課題が存在し、製品や工程の種類によって有効なアプローチが異なる可能性がある。例えば、工程間で製品の構造に大きな変化がある場合は局所の特徴が有効である一方で、外観変化が乏しい工程では大域の特徴の方が効果的に機能すると考えられる。従って、製品や工程の条件に応じて最適なアプローチを選択・融合できるように、両手法の定量的な比較検討が必要不可欠である。さらに、実際の工場環境では学習データの取得が困難であることが多く、限られたデータ量でも高精度に進捗推定が可能な手法の検討も現場実装に向けた重要な要件である。

本研究では、部品検出に基づく局所の特徴アプローチと、深層距離学習に基づく大域の特徴アプローチの2つの手法を用いて、同一のデータセットにおいて比較検討を行う。各手法の精度、工程間の誤認識傾向、オクル

ージョンへの耐性、少量データでの汎化性能といった観点から比較し、今後の融合モデル構築に向けた有効性と課題を明らかにすることを目的とする。また、部品検出手法として従来 [2] 使用していた YOLACT [4] から YOLOv8 [5] への変更を行う。これは深層距離学習手法において YOLOv8 を用いているため、両手法で共通の検出モデルを用いることで、将来的な融合モデルの構築において処理系の統一と連携が容易になると考えている。

2. 比較検討する手法

2.1 インスタンスセグメンテーションでの部品検出に基づいた進捗推定手法 [2]

湯本らは、YOLACT を用いたインスタンスセグメンテーションによって部品を検出し、検出された部品の種類と数に基づいて組立進捗を推定する手法を提案した [2]。本手法は、進捗ステップごとに取り付けられる部品の組み合わせを定義し、推論時に検出された部品と照合することで、該当する製品の組立進捗を判断する。つまり該当しないパターンをエラーと判定するため、学習時にエラーデータを含めなくてよい。さらに、アノテーションデータを利用して、部品が完全に隠れないように制御された矩形マスクを複数重畳するデータ拡張を行い、学習データの多様化と検出精度の向上を実現した。

2.2 Anomaly Quadruplet-Net での進捗推定手法 [3]

三浦らは、組立工程間で外観の変化が小さい作業に対しても高精度な進捗推定を可能にするために、Anomaly Quadruplet-Net [3] を提案した。本手法は、従来の Anomaly Triplet-Net が3サンプル間の相対距離 (anchor, positive, negative) に基づいていたに対し、4サンプル (anchor, positive, negative1, negative2) を用いた Quadruplet [6] を導入することで、隣接ステップ間の誤判定をより効果的に抑制できるよう設計されている。さらに、オクルージョンなどの実環境下での欠損を模擬した Anomaly サンプルを Random Erasing [7] で生成し、距離空間上で正常サンプル群から十分に離れるよう学習する損失関数を追加している。これにより、外観変化が微小な工程間や遮蔽が生じる環境においても進捗推定が可能となり、Triplet ベースの従来手法と比較して精度と誤判定率の両面で改善が確認された。

3. 比較実験

3.1 実験概要

工場の製品に模したデータセットとして、デスクトップパソコンとカラーボックスの二つの対象製品のデータセットを用意した。デスクトップパソコンは部品点数が多く、工程の視覚的識別が困難な対象であり、カラーボックスは段階的かつ明確な進捗変化を示す比較的単純な対象である。加えて、撮影環境や背景の複雑さの違いも含まれており、各手法の環境変動に対する頑健性も併せて検証できる設計とした。組立工程は、部品が1種類が増えるとステップも1つ増えるように定義している。デスクトップパソコンでは、部品全8種類 (case, power, board, CPU, fan, GPU, HDD, cover) とし、計8ステップとなるように定義している。またカラーボックスでは、部品全7種類 (frame, board, basket, book, gel bottle, tumbler, cover) とし、計7ステップとなるように定義している。深層距離学習手法では、各対象においてオクルージョン画像を新たなクラスとして追加した。図1、図2に各ステップの画像の例を示す。デスクトップパソコンの各ステップの枚数は、trainで40枚、validationで150枚、testでは129枚とした。またカラーボックスの各ステップごとの枚数は、trainで40枚、validationで150枚、testでは150枚とした。

本実験では、部品検出手法と Anomaly Quadruplet-Net による深層距離学習手法の2つを評価するために、trainとval画像を用いた学習性能評価、test画像に対する進捗推定精度の算出、作業動画に対する性能評価の3つを行った。



図1 デスクトップパソコンの進捗画像

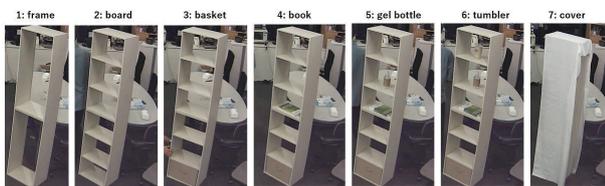


図2 カラーボックスの進捗画像

3.1.1 学習性能評価

2つの進捗推定手法を用いて、それぞれのデータに対して、エポック数を100、バッチサイズを4、学習率を0.0001の同条件で学習させ、性能評価として学習時間と損失関数を比較する。学習時間の比較を表1、損失関数の比較を図3、図4に示す。

本評価では、深層距離学習手法は、部品検出手法に比べて学習時間を要するものの、特にカラーボックスにおいてtrainとvalの損失差が小さく、安定した学習が確認された。一方部品検出手法では後半にかけて損失差が拡大し、過学習の傾向が顕著となった。深層距離学習手法は、学習時間は多く要したが、部品検出手法と異なりアノテーションコストを必要としない。以上の結果から、学習の安定性の観点においては、深層距離学習手法の方が有効であると考えられる。

表1 各手法における学習時間の比較 (単位: 時間)

手法	デスクトップパソコン	カラーボックス
深層距離学習	10.11	8.62
部品検出	1.27	0.91

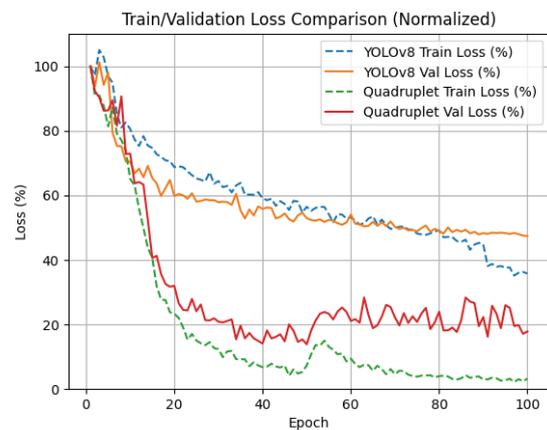


図3 デスクトップパソコンの各損失関数のグラフ

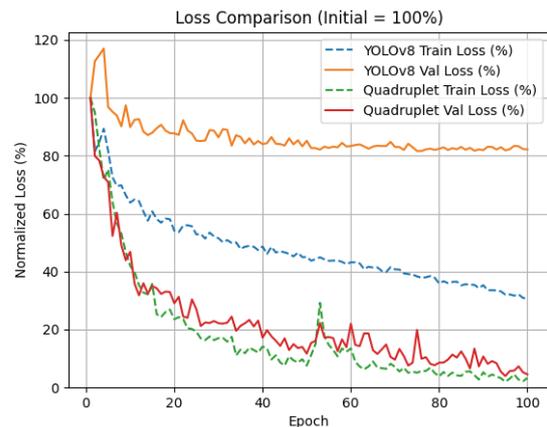


図4 カラーボックスの各損失関数のグラフ

3.1.2 test画像に対する進捗推定精度の算出

次に、2つの手法の進捗推定精度の比較を行うため、学習したモデルを用いて2つのデータセットにおけるtest画像で推定を行った。進捗推定精度は各データセットに対して混同行列から算出した。各手法の2つのデータセ

ットにおける混同行列を図5, 6, 7, 8に示す. また各混同行列の進捗推定精度を表2に示す.

デスクトップパソコンデータでは, 深層距離学習手法において高い進捗推定精度が得られた. 特にエラー判定に関しては, 部品検出手法の誤判定先がステップ1と偏りが生じることにに対し, 深層距離学習手法では高い精度が見られた. これは, 深層距離学習手法では多様な異常パターンに対応していると考えられる.

一方, カラーボックスデータでは, 背景が見えることによる変動に対するモデルの頑健性が問われる状況となったが, 部品検出手法が高い分類精度を示した. 全ステップでの分類がほぼ完全であり, 特に中間工程であるステップ5およびステップ6において, 深層距離学習よりも安定した精度を維持していた. さらに, 誤判定クラスにおいても部品検出手法が深層距離学習手法より高い精度を示した.

表2 各データにおける進捗推定精度 (単位: %)

手法	デスクトップパソコン	カラーボックス
深層距離学習	96.0	87.1
部品検出	94.3	93.8

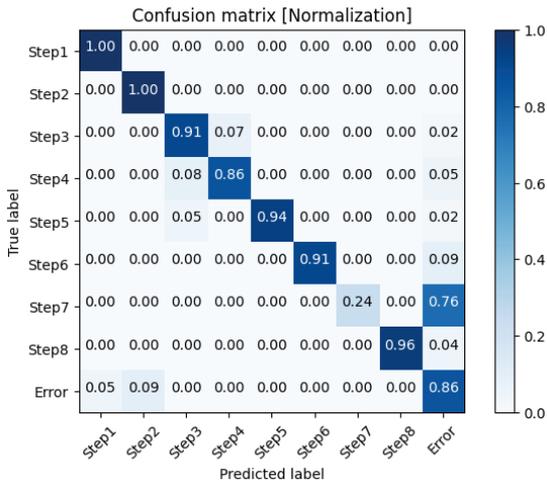


図5 深層距離学習手法によるデスクトップパソコン

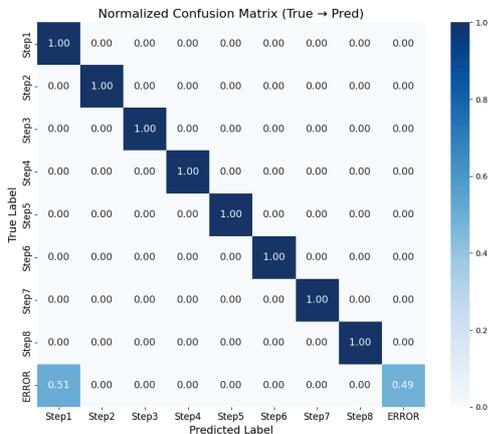


図6 部品検出手法によるデスクトップパソコン

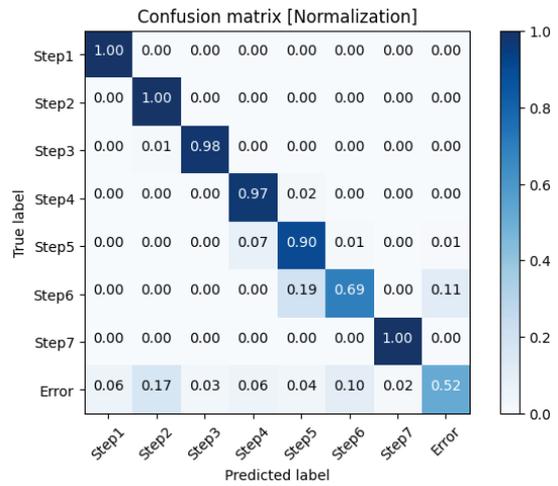


図7 深層距離学習手法によるカラーボックス

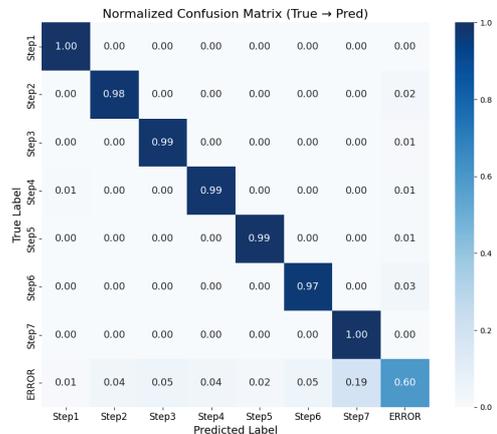


図8 部品検出手法によるカラーボックス

3.1.3 作業動画に対する性能評価

デスクトップパソコンおよびカラーボックスの組立作業動画に対し, 深層距離学習手法と部品検出手法による進捗推定を比較した. 工程ごとの進捗を連続で推定するためには, 同一製品の識別と追跡が不可欠であることから, YOLOv8による物体検出と SORT [8]によるトラッキングを行った. 推定されたステップと時間(フレーム)との関係を図9, 10, 11, 12に示す. 図中では, 黒色の印を各フレームで推定されたステップ, 緑色の印を5フレーム連続で同一ステップが推定された場合の確定ステップとして示す. 加えて, 赤色の印は推定結果が「ERROR (step = -1)」となったフレームを表しており, エラー判定を可視化している.

図10と図12より, 部品検出手法は全体的に段階的かつ安定したステップ推定が可能であり, 短時間の作業変化や終盤工程においても推定していることを示した. 一方深層距離学習手法では, 図9と図11より, 中間ステップの識別誤りや最終ステップ未検出などの課題が見られた. これは, 部品検出手法が個別部品の有無を基にした推定により環境変動に強いのにに対し, 深層距離学習手法は画像全体の外観に依存するため微小な外観変化や遮蔽に対して脆弱であることが要因と考えられる.

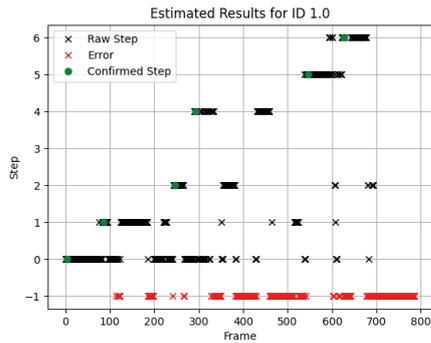


図9 深層距離学習手法のデスクトップパソコンにおける推定ステップと時間の関係

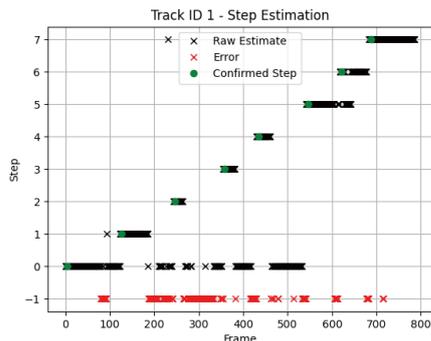


図10 部品検出手法のデスクトップパソコンにおける推定ステップと時間の関係

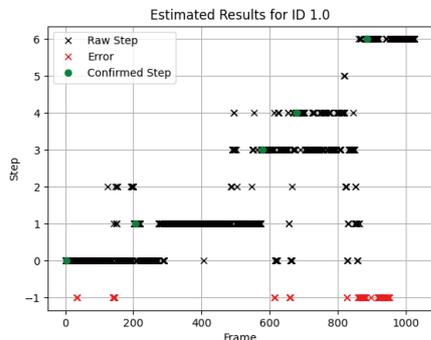


図11 深層距離学習手法のカラーボックスにおける推定ステップと時間の関係

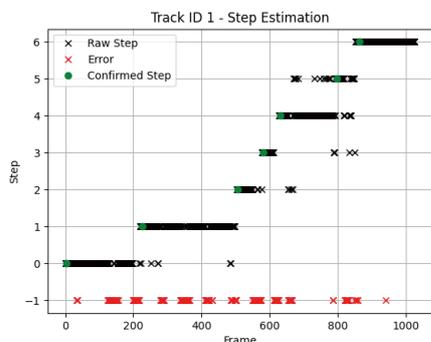


図12 部品検出手法のカラーボックスにおける推定ステップと時間の関係

4. 結論

本研究では、部品検出手法と深層距離学習手法の2つのアプローチに対して、組立進捗推定における特性と課題を定量的に比較検討した。部品検出手法は、作業中の遮蔽がある環境下でも、各部品の有無を明確に捉えることで、高い分類精度と段階的な進捗推定を実現した。一方、深層距離学習手法は、外観変化の小さい工程で誤判定が生じる傾向が見られたものの、学習時の安定性や異常検知性能において一定の有効性が確認された。

これらの結果から、各フレームにおいて部品検出手法により認識された部品の有無や種類をステップ判定の主要指標としつつ、深層距離学習手法によって算出された画像全体の特徴ベクトルを補助的な判断要素として組み込むことで部品が一時的に遮蔽された場合や判別が困難な中間工程においても、より信頼性の高い推定を可能とするようなハイブリッド設計を今後検討する。

謝辞

本研究の実施にあたり、株式会社日立ハイテクの新見滉希氏他にご助言頂いた。記して謝する。

参考文献

- [1] J. Wang, Y. Ma, L. Zhang, 他 :“Deep learning for smart manufacturing: Methods and applications”, Journal of Manufacturing Systems, Vol. 48, pp. 144–156, 2018.
- [2] 湯木茂樹, 三浦一真, Alessandro Moro, Sarthak Pathak, 梅田和昇 :“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”, 動的画像処理実利用化ワークショップ DIA2024 講演論文集, IS1-16, pp. 73–77, 2024.
- [3] 三浦一真, Sarthak Pathak, 梅田和昇 :“外観特徴の変化が小さい作業を考慮した深層距離学習による組立作業の進捗推定”, 動的画像処理実利用化ワークショップ DIA2025, IS3-25, 2025.
- [4] D. Bolya, C. Zhou, F. Xiao, Y. J. Lee :“YOLACT: Real-time instance segmentation”, Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV), pp. 9157–9166, 2019.
- [5] R. Varghese, S. M. :“YOLOv8: A Novel Object Detection Algorithm with Enhanced Performance and Robustness”, 2024 International Conference on Advances in Data Engineering and Intelligent Computing Systems (ADICS), pp. 1–6, India, 2024.
- [6] W. Chen, X. Chen, J. Zhang, 他 :“Beyond Triplet Loss: A Deep Quadruplet Network for Person Re-identification”, 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 1320–1329, 2017.
- [7] Z. Zhong, L. Zheng, G. Kang, S. Li, Y. Yang :“Random erasing data augmentation”, Proceedings of the AAAI Conference on Artificial Intelligence, Vol. 34, No. 07, pp. 13001–13008, 2020.
- [8] A. Bewley, Z. Ge, L. Ott, F. Ramos, B. Upcroft :“Simple Online and Realtime Tracking”, Proceedings of the IEEE International Conference on Image Processing (ICIP), pp. 3464–3468, 2016.