

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

○三浦一真 †, Sarthak Pathak ‡, 梅田和昇 ‡

○ Kazuma MIURA † Sarthak PATHAK ‡ and Kazunori UMEDA ‡

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

‡ 中央大学理工学部精密機械工学科,

＜要約＞本研究では、外観特徴の変化が小さい作業を考慮した、工場内での製品の組立進捗判定を行う手法を提案する。工場での実用を考慮して、少ない学習データ数で使用可能な深層距離学習を行う。Quadruplet Loss を利用することで、組立作業による製品の 외観特徴の変化が小さい場合でも隣接する作業の誤認識を減少させることが期待される。デスクトップパソコンの組立作業で実験を行い、手法の有効性を確認した。

＜キーワード＞深層距離学習, 進捗推定, 組立作業, スマートファクトリー

1 序論

近年、工場のスマートファクトリー化が進められている [1]。しかし、人手による作業は機械による作業に比べて場所や手順が均一でないことからセンシングが難しく、作業の進捗管理を人の手で紙に記すなどの方法で行っているという課題が存在する。この課題を解決するために、作業者の動作認識を行い進捗推定を行う研究 [2][3] が行われている。一方で、実際の工場では組立手順が明確に定まっていない製品や、1つの作業に数時間かかる製品も組み立てられており、作業者の動作認識を行うことが困難な場合が存在する。そこで、作業者の動作ではなく、製品の 외観特徴に着目した研究も行われている。例として、製品の 외観特徴のうち部品に着目した研究として、インスタンスセグメンテーションを用いて部品を認識することで、組立作業の進捗を推定する研究 [4] が行われている。しかし、工場のようなオクルージョンの多い環境下では、検出部品が小さい場合に部品の検出精度が低下してしまう。部品ではなく製品全体の 외観特徴に着目した研究として、深層距離学習を用いて組立作業の進捗を推定する研究 [5] が行われている。組立進捗やオクルージョンに対し、深層距離学習を行うことで、クラス分類問題として、オクルージョンも考慮した組立作業の進捗推定に成功している。しかし、連続して行われる 외観特徴

の変化の小さい作業間では誤認識が生じてしまう。そこで本研究では、組立作業を対象の製品の 외観特徴に着目した、オクルージョン及び外観特徴の変化が小さい作業に対して頑健なシステムを構築することを目的とする。

2 従来手法: Anomaly Triplet-Net

Anomaly Triplet-Net[5] は、製品の進捗状況を Step で認識してクラスとして認識を行うことで、Triplet Loss[6] を用いた深層距離学習を行っている。人が見ても進捗状況がわからないオクルージョンの存在する製品画像を進捗推定可能なクラスから離すように学習することで、実環境でも精度が期待できる手法である。さらに、Anomaly Triplet-Net によって得られた推定結果に誤認識が生じた際、組立作業動画の前フレームでの進捗推定と大きく異なっていれば、進捗推定結果が連続的になるように補正することで誤認識を減少させる。しかし、前フレームでの進捗推定から連続的に行われる作業として誤って推定された場合、補正によって誤認識を減少させることができない。このように連続的に行われる作業として誤認識されてしまった場合は、進捗推定精度に重大な悪影響を及ぼしてしまう。

3 提案手法

3.1 提案手法概要

提案するシステムの概要図を図 1 に示す。進捗推定を行う提案モデルには、Quadruplet[7] を利用する。Quadruplet を利用するにあたり、距離学習を行う際のクラス及びデータ選択の組み合わせは爆発的な増加を示す。よって、効率的に学習が進むようデータローダを工夫する。

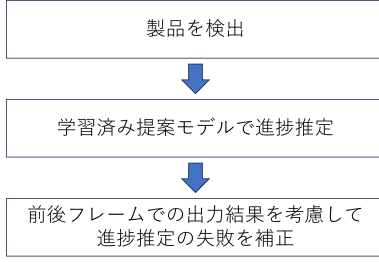


図 1 システムの概要

3.2 進捗推定モデル

連続的に行われる作業間では、進捗状況が進んでも製品の外観特徴の変化は小さい場合が存在する。そこで、隣接する作業を選択して距離学習を行う事で、誤認識を減少させる。また、提案する Anomaly Quadruplet-Net は、Anomaly Triplet-Net[5] にインスパイアされており、特徴量空間上でオクルージョン画像を Quadruplet で学習されるデータから離すように距離学習を行うことで、組立作業の開始から終了までの誤認識を減少させることが期待される。学習時には anchor, positive, negative1, negative2 が選択され、anomaly は Random Erasing によって生成する。選択された画像は CNN によって特徴量抽出を行い、最終的に 64 次元の特徴量空間に受け渡される。その後、特徴量空間上で距離学習を行う。推定時には k-NN 用いて、推定対象のデータが属する進捗状況の判定を行う。

3.3 損失関数

損失関数は、anchor, positive, negative1, negative2 サンプルの 4 つを 1 組として相対的距離を算出する。Quadruplet Loss と、オクルージョンが存在するサンプルを anomaly として相対的距離を算出する Anomaly Triplet-Net を組み合わせている。提案する損失関数を式 (1) に示す。

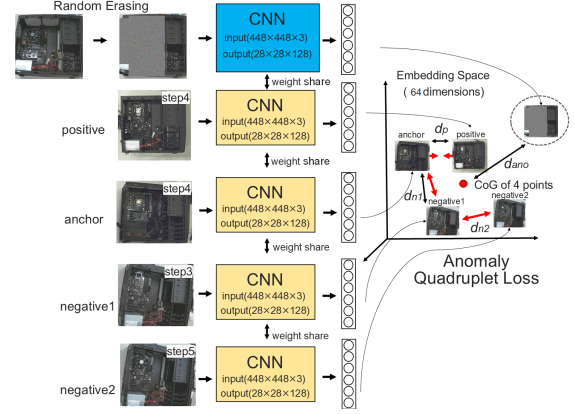


図 2 Anomaly Quadruplet-Net の概要

$$\begin{aligned}
 L_{AnomalyQuadruplet} = & \max(d_p - d_{n1} + m_\alpha, 0) \\
 & + \max(d_p - d_{n2} + m_\beta, 0) \\
 & + \lambda \cdot \max(d_p - d_{ano} + m_c, 0) \quad (1)
 \end{aligned}$$

d_p, d_{n1}, d_{n2} , はそれぞれ、positive サンプル, negative1 サンプル, negative2 サンプルの anchor サンプルからの特徴量空間上の距離を表す。 d_{ano} は anomaly サンプルと anchor サンプル, positive サンプル, negative1 サンプル, negative2 サンプルの 4 点の重心との特徴量空間上の距離を表す。 m_α, m_β, m_c は学習を行う際に動かすサンプル間の距離を示すマージンである。 λ は anomaly サンプルの学習を始める epoch を調整する重みである。特徴量空間上の距離はユークリッド距離を用いる。

3.4 データローダ

データセットが n 個のクラスを持ち、1 クラスあたり m 個の画像があるとき、Triplet Loss でのデータの組み合わせ数は式 (2)、Quadruplet Loss でのデータの組み合わせ数は式 (3) で表される。

$${}_nC_1 \times {}_mP_2 \times {}_{n-1}C_1 \times {}_mP_1 \quad (2)$$

$${}_nC_1 \times {}_mP_2 \times {}_{n-1}C_1 \times {}_mP_1 \times {}_{n-2}C_1 \times {}_mP_1 \quad (3)$$

例として、クラスが 8 個、各クラス 40 枚の画像のデータセットに適応すると、Triplet Loss では約 350 万通り、Quadruplet Loss では約 8 億通りのデータセットの組み合わせ数となる。これは計算量が膨大になり処理時間や処理能力が問題である。そこで本手法では、連

続するクラスは連続する作業進捗を表していることから、外見特徴が類似していることに着目して、連続する3クラスの組み合わせのみを選択することで、学習するデータの組み合わせ数を減少させるよう工夫する。Anomaly Quadruplet-Net では、anchor サンプルを n 個のクラスに対し1枚ずつ選択し、positive サンプルは anchor サンプルと同じクラスかつ別の画像を選択する。negative1 サンプル, negative2 サンプルは、anchor サンプルのクラスより一つ小さいクラスと一つ大きいクラスから選択する。anomaly サンプルは anchor サンプル以外から選んだ画像に対して、Random Erasing を行い生成する。類似クラスを十分に離すように学習するため、negative1 サンプル, negative2 サンプルのクラスを入れ替えて再度組み合わせを選択する。よって、 $2 \times n \times m$ 通りとする。図3に学習のイメージ図を示す。

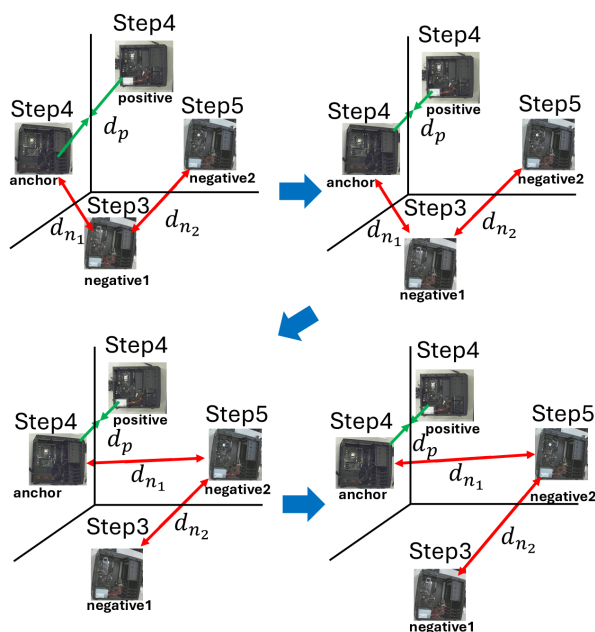


図3 学習の組み合わせ

4 実験

4.1 実験概要

デスクトップパソコンの組立作業を8クラスの進捗に分割し、実験を行った。図4に各Stepの画像の例を示す。各ステップの枚数は、Trainで各40枚の計320枚、Validarionで各150枚の計1200枚、Testではオクルージョン画像を9クラス目と追加し、各129枚の計1161枚とした。Test画像に対する進捗推定精度を算出

し、Anomaly Quadruplet-Net を評価した。評価を行う為、従来手法である Anomaly Triplet-Net と提案する進捗推定モデルである Anomaly Quadruplet-Net の二つのモデルで実験を行う。

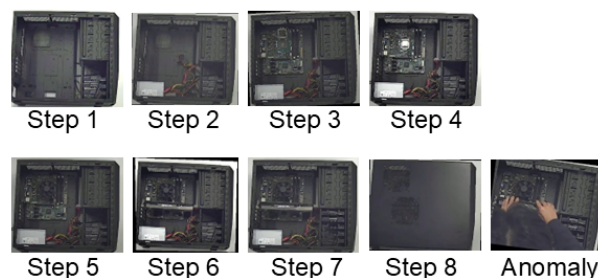


図4 デスクトップパソコンデータセットの例

4.2 Anomaly Triplet-Net での実験

Anomaly Triplet-Net での実験条件は経験則から決定した。提案した Anomaly Quadruplet と条件を近づける為に、特徴量空間は32次元に変更して実験を行った。また、学習時の Epoch 数は合計で200Epochとし、Anomaly 画像の学習は50Epochから行うように設定し学習を行った。図5に、混同行列を示す。Anomaly 画像も含めた Test データ全体に対する精度、すなわち Error 判定も含めた場合の精度は96.5%となった。また、隣接 Step での誤判定率は4.8%となった。図6に、学習済みの特徴量を T-SNE[8] で2次元で可視化した結果を示す。

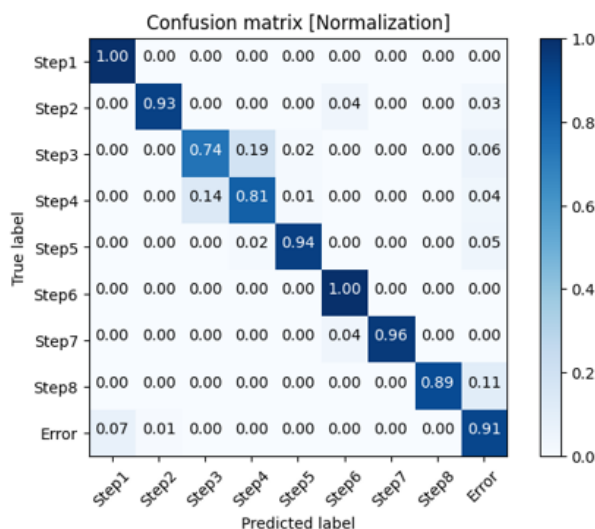


図5 Anomaly Triplet での進捗推定

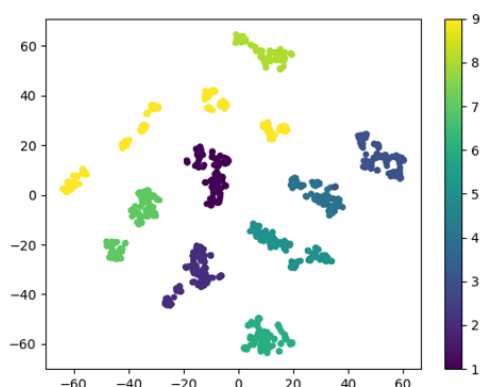


図 6 Anomaly Triplet 学習結果の可視化

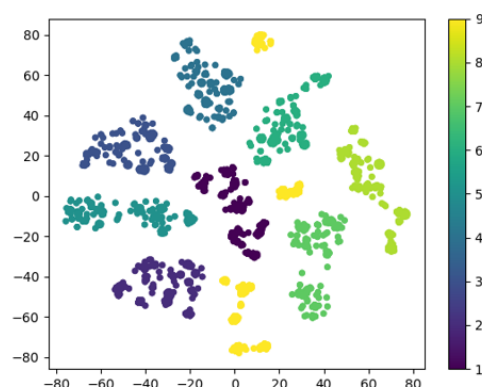


図 8 Anomaly Quadruplet 学習結果の可視化

4.3 Anomaly Quadruplet-Net での実験

Anomaly Quadruplet-Net での実験条件は経験則から決定した。特徴量空間は 64 次元，合計 250Epoch，Anomaly 画像の学習は 100Epoch から行った。図 7 に，混同行列を示す。Anomaly 画像も含めた Test データ全体に対する精度は，97.8%，隣接 Step での誤判定率は 2.9% となった。図 8 に，学習済みの特徴量を T-SNE で 2 次元で可視化した結果を示す。

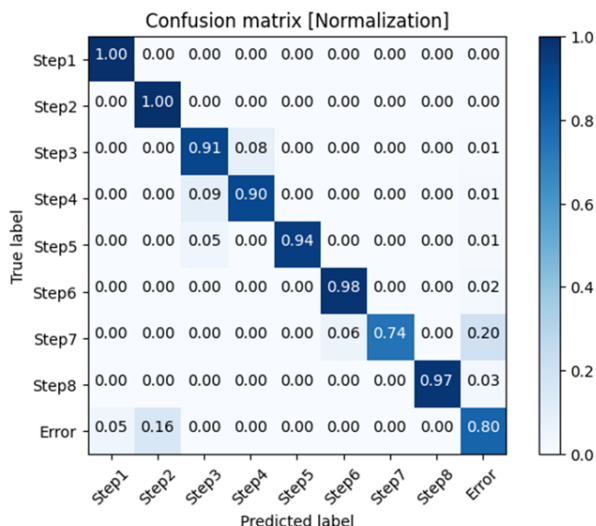


図 7 Anomaly Quadruplet での進捗推定

4.4 考察

デスクトップパソコンデータセットを使用した実験の結果から，提案した Anomaly Quadruplet-Net は従来手法に対して，Accuracy が 1.3% 向上することがわかった。また，隣接作業での誤判定を 1.9% 減少させる事がわかった。特に図 5 と 7 の混同行列の結果から，Step3

と Step4 の間での誤判定を減少させる事に成功していることがわかる。これは，マザーボードに CPU を取り付ける前と後の見た目の変化が小さい作業である。よって，本手法で提案した外観特徴の変化が小さい作業に対して頑健にする為に提案された Loss および，データローダが手法の有効性が確認できた。一方で，図 6 と 8 から，両手法とも Anomaly の判定自体は成功していることがわかる。しかし，Anomaly の判定精度に関しては，提案手法は従来手法より精度が低下している事がわかる。これは，Anomaly と Anomaly 以外のサンプルを離しながら学習する際に，従来手法では，anchor, positive, negative の 3 点重心と anomaly との距離を離すような Loss を設計していたが，本手法では，anchor, positive, negative1, negative2 の 4 点重心との距離に変更している為，Anomaly の学習が従来手法に比べて上手く行かなくなった可能性が考えられる。しかしながら，動画に対して本手法を適応する際には，進捗推定結果を時系列処理を行い補正を行う想定であるため，Anomaly の判定精度の低下は後処理で十分に対応可能なものであると考えられる。

5 結論

本研究では，外観特徴の変化が小さい作業を考慮した，Anomaly 画像と Quadruplet Loss を利用した組立作業の進捗推定手法を提案した。実験では，連続する作業に対して深層距離学習を行うことで，97.8% の精度で進捗推定を行える事を確認した。また，隣接 Step での誤判定率は 2.9% に抑えることに成功した。今後は作業動画に対して Anomaly Quadruplet-Net を適応し，

動画から得られた進捗推定結果を時系列的に補正することで誤認識を減少させる。

参考文献

- [1] Wang, J., Ma, Y., Zhang, L., et al.: “Deep learning for smart manufacturing: Methods and applications”, Journal of Manufacturing Systems, Vol.48, pp.144-156 (2018)
- [2] 日立産業制御ソリューションズ: “AIによる人物姿勢・動作認識ソリューション”, Available from:https://info.hitachi-ics.co.jp/product/activity_evaluation/
- [3] 夏 清心: “Research on Sensor Data Processing Methods for Recognizing and Understanding Factory Work Activities”, 研究報告ユビキタスコンピューティングシステム (UBI), Vol.31, pp.1-2 (2022)
- [4] 湯本茂樹, 三浦一真, 梅田和昇, et al.: “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 2024 講演論文集, pp.73-77 (2024)
- [5] 橘川拓実, 湯本茂樹, 梅田和昇, et al.: “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 2023 講演論文集, pp.107-113 (2023)
- [6] Hoffer, E., and Ailon, N.: “Deep metric learning using Triplet network”, International Workshop on Similarity-Based Pattern Recognition, pp.84-92 (2015)
- [7] Chen, W., Chen, X., Zhang, J., et al.: “Beyond Triplet Loss: A Deep Quadruplet Network for Person Re-identification”, 2017 IEEE Conference on Computer Vision and Pattern Recognition, pp.1320-1329 (2017)
- [8] Maaten, L., and Hinton, G.E.: “Visualizing Data using t-SNE”, Journal of Machine Learning Research, Vol.9, pp.2579-2605 (2008)