

Deep Metric Learning を用いた組立作業の進捗判定

Progress Estimation of Assembly Work Using Deep Metric Learning

中央大学 ○橋川 拓実, Alessandro Moro, Sarthak Pathak, 梅田 和昇
日立ハイテクソリューションズ 原田 佳周, 西川 英雄, 野口 稔, 濱谷 章史

Chuo University ○Takumi Kitsukawa, Alessandro Moro, Sarthak Pathak, Kazunori Umeda
Hitachi High-Tech Solutions Corporation, Yoshihiro Harada, Hideo Nishikawa, Minoru Noguchi, Akifumi Hamaya

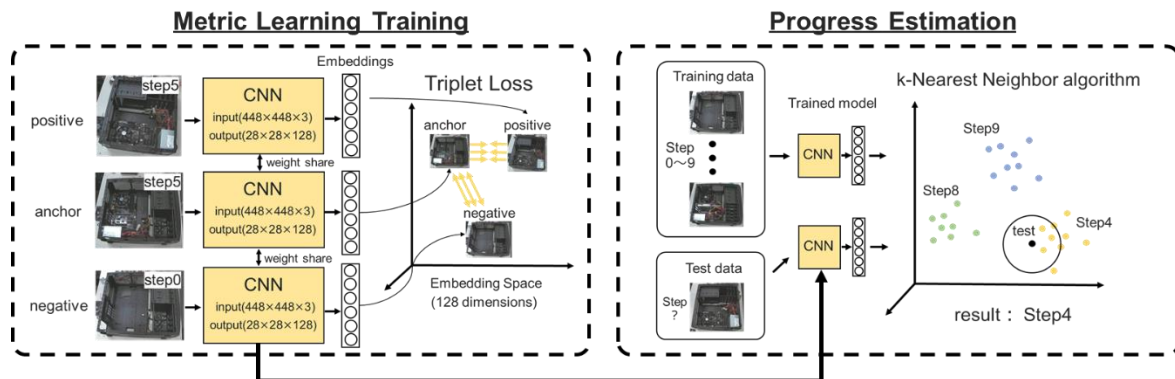


Fig. 1 構築したモデルのネットワーク構造

1. 序論

近年、製造業における人手不足は深刻度を増しており、現場の作業効率を上げて生産性を向上させることが求められている。解決方法として IoT やロボットを利用した製造ラインの自動化が進められている。しかし、顧客のニーズに合わせて製品を生産する多品種少量生産型の工場では、自動化が進んでおらず、人手による作業が大半を占めている。特に製品の組立作業は、商品の仕様が変わるたびに生産ラインを変更する必要があるため、人手による作業が行われている場合が多く、作業の進捗管理を紙などのアナログな方法で行っているという問題がある。これに対し、工場に設置した定点カメラから映像を取得し作業の進捗状況を推定することで、自動で管理するシステムの導入が有効であると考えられる。これまでに、人物の骨格点情報を時系列的に利用し動作認識する研究[1]や作業者目線で手の動きを見ることで進捗推定する研究[2]が行われている。

しかし大型装置の組立は、作業が数日かかるものや組立順序が細かく決まっていないものも多く、作業者の動きや時系列データを利用した進捗推定が適していない。

そこで本研究では、組立てる対象物に着目し、組立作業の進捗を推定するシステムを構築することを目的とする。また進捗推定は、先に設定した組立進捗ステップに対し、どのステップまで作業が進んでいるのかを判定する方法を用いる。

2. Metric Learning の利用

設定したステップに対し現在どのステップに属するかを判定するシステムの構築としては、クラス分類問題として問題設定することができると考えられる。画像におけるクラス分類は、CNN (Convolutional Neural Network)を用いた深層学習手法が高い精度を誇っている。しかし今回の組立進捗推定におけるクラス分類は、クラス同士の見た目の違いが小さいことが考えられ、細かな特徴の違いを抽出することが求められる。そこで機械学習手法の一つである Metric Learning (距離学習)[3]の利用を検討する。

Metric Learning は、サンプル間の距離または類似度に基づいてクラスごとに分離されるように、入力データから特徴量空間へ変換するパターンを学習させる手法である。Metric Learning の目的は、同じクラスのサンプル間の距離を小さくしながら異なるクラスのサンプル間の距離を大きくすることである。特にパターンの学習の際に多層構造のニューラルネットワークを用いて特徴抽出をする Deep Metric Learning (深層距離学習)[4]も提案されている。

一般的なクラス分類は CNN で抽出した特徴量を全結合層に通し、それをソフトマックス関数でクラス所属確率に変換する。し

かしこれだと同クラスと他クラスのサンプル間の距離を考慮せずに特徴量抽出ネットワークを学習させることになる。一方 Deep Metric Learning は、意図的に他クラスのサンプル間を大きくし、同クラスのサンプル間の距離を小さくするように学習することで識別的な特徴量を得ることができる。そのため各クラスのサンプルが少ないときや未知クラスがあるときに性能を発揮する。

今回の組立作業の進捗判定は、実際の工場での実用化を考えると、少ないデータでの学習が求められる。また、未知クラス判定により誤判定を抑えられることから、本研究では Deep Metric Learning を利用したステップ判定手法 (Fig.1 参照)を提案する。

3. 提案手法

3.1 システムの概要

提案するシステムの概要を Fig.2 に示す。事前の準備として、画像内での対象物の位置を検出できるようにするために、カスタムデータセットによる物体検出モデルの学習を行う。また、切り出し画像の進捗推定を行うステップ判定モデルを学習させる。

システムの流れとしては、定点カメラから取得した画像に対し物体検出手法で対象物を検出、その位置を切り出し Deep Metric Learning によるステップ判定モデルで進捗を推定する。物体検出手法は、高速でありながら正確な検出が可能な Faster R-CNN[5]の学習済みモデルをファインチューニングして利用する。

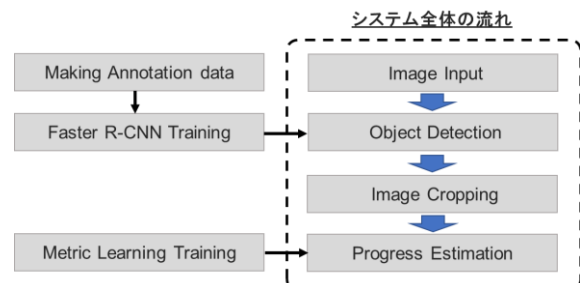


Fig. 2 提案システムの概要

3.2 ステップ判定手法

ステップ判定は、Fig.1 の手順に従って行う。まずは組立進捗の中で判定したいステップを設定して学習データを用意する。方法は、組立動画から対象物部分を切り出しそれぞれのステップごとに分けて保存する。その学習データの中からランダムに画像を一枚選び、anchor サンプルを設定する。この anchor サンプルと同ス

ステップの画像を一枚選んだものを **positive** サンプル、他ステップの画像を一枚選んだものを **negative** サンプルとする。次にこれらの設定した三枚の画像をそれぞれ 4 層の CNN モデルに入力、さらに 1 層の全結合層に入力することで 128 次元の特徴ベクトルを取得する。この時 CNN モデルの重みは共有する。そして得られた三つの特徴ベクトルのうち、**anchor** サンプルと **positive** サンプルの距離を近づけ、**anchor** サンプルと **negative** サンプルの距離を遠ざけるように損失関数を定義することでパラメータの更新を行う。損失関数には **anchor** サンプル、**positive** サンプル、**negative** サンプルの三枚を一組のデータとして、相対的距離を算出する **Triplet Loss**[6]を用いる。式は以下である。

$$L_{\text{triplet}} = \max(d_p - d_n + m, 0)$$

d_p, d_n は、特徴量空間内における **positive** サンプル、**negative** サンプルと **anchor** サンプルのそれぞれの距離である。 m は、距離を遠ざける・近づける作業の動かす程度を表す任意定数であり、マージンと呼ぶ。距離の計算にはユークリッド距離を利用する。

次に、学習した CNN モデルを利用してステップ判定を行う。まず学習で利用した各ステップの画像をもう一度一枚ずつ学習済みモデルに入力し、特徴量空間に埋め込む。すべての画像を特徴量ベクトルに変換したら、推定したい入力画像を同じく学習済みモデルに入力し、特徴量空間に埋め込む。これらの埋め込まれたデータに対し k 近傍法 (k -Nearest Neighbor algorithm) を利用することで未知のテストデータがどのステップに属するかを推定する。さらに判定の際、特徴量空間内での距離がしきい値より遠いものをエラーと判定することで、組立作業における組立ミスの検知やオクルージョンで生じる誤判定の削減をすることができる。

また組立作業における進捗ステップの判定は、隣のステップとの特徴が似ていることから、近いステップとの誤判定が起きやすいと考えられる。そこで、利用する **Triplet Loss** のマージンの値を近くステップの時に大きくなるように適応的に変化させる手法を提案する。計算式は近くのものから遠いものまで滑らかにマージンが下がるように正規分布を利用する。提案する適応的マージンの式は以下である。

$$m = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(n_n - n_a)^2}{2\sigma^2}\right) \cdot a$$

σ^2 はステップ数の分散、 a はマージンの大きさを表す任意定数、 n_a は **anchor** サンプルのステップ数、 n_n は **negative** サンプルのステップ数である。

4. 実験

4.1 ステップ判定に対する実験

ステップ判定モデルの有効性を確認するために実験を行った。組立製品の対象としてデスクトップ PC を使用した。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。判定するステップは、ステップ 0 から 9 の 10 個のステップに設定した。ステップの分け方は Fig.3 に示す。

実験は **Triplet Loss** のマージンを定数にした手法と適応的マージンを利用した手法で比較した。学習に用いた画像は各ステップ 4 枚の計 40 枚で、 k 近傍法の定数 k は $k=4$ 、学習は 10000epoch 行った。また学習データが少量であるためデータに偏りが生まれないように、入力画像に対して回転、射影変換、色彩変化、一部切り落としをランダムに加えることでデータを拡張した。実験結果を Fig.4 に示す。

マージン固定では 10000epoch で 82.0[%]、適応的マージンでは 91.8[%]の正解率を達成することができた。また混同行列から、どちらの手法でも高い精度で判別に成功していることが確認できた。

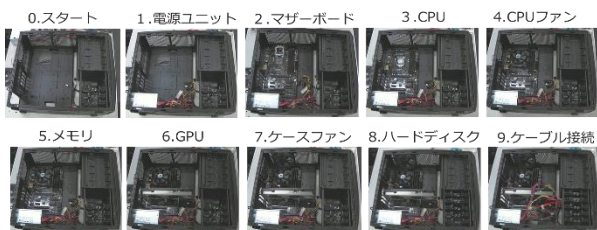


Fig. 3 進捗ステップの分け方

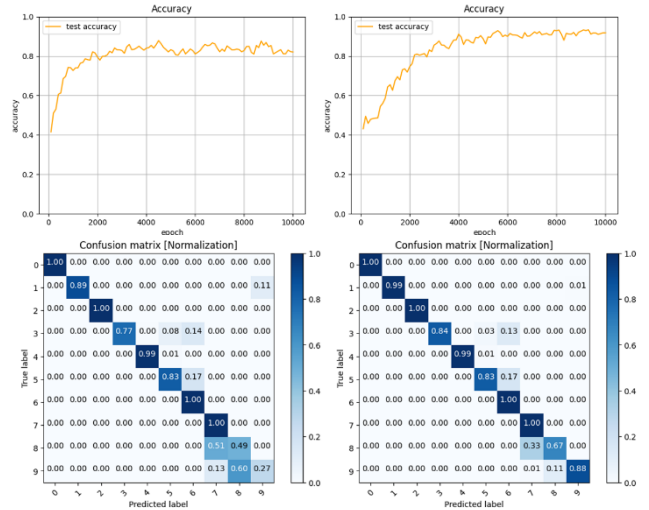


Fig. 4 実験結果：上段からAccuracy, 混同行列、左列からマージン固定、適応的マージン

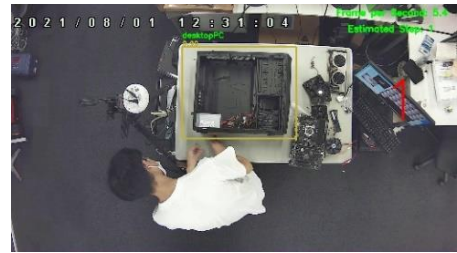


Fig. 5 システム全体の実験結果

特に適応的マージンを用いた手法では、隣のステップとの誤判定が減少しており、有効性が検証できた。

4.2 システム全体に対する実験

物体検出から、切り出し、ステップ判定の一連の流れに対する実験を行った。物体検出においては約 1000 枚の画像に対しラベル付けを行い、Faster R-CNN をファインチューニングした。また作業員によるオクルージョンで生じる誤判定を避けるため、20 フレーム連続で同じ判定が出たら判定ステップを固定し、ログとして残す方法を利用した。

実験の結果、すべてのステップで正しい判定をログとして残すことができた。実際のシステムを動かした様子を Fig.5 に示す。黄色のパウンディングボックスが対象物の検出を表しており、右側の赤文字の数字が判定されたステップを表している。これらの結果から、物体検出を含めたシステム全体の有効性も検証できた。

5. 結論

本研究では、対象物に着目した組立作業の進捗推定システムの構築を行った。ステップ判定には **Deep Metric Learning** を用いた推定手法を提案し、実験でその有効性を検証した。

今後の展望として、検出した対象物の時系列情報を記録するために、トラッキング手法を取り入れたシステムの構築を目指す。

参考文献

- [1] 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", https://info.hitachi-ics.co.jp/product/activity_evaluation/, 2022-07-08.
- [2] M. Funk, A. Bachler, L. Bachler, T. Kosch, T. Heidenreich, A. Schmidt, "Working with Augmented Reality? A Long-Term Analysis of In-Situ Instructions at the Assembly Workplace," PETRA'17, 2017.
- [3] E. Xing, M. Jordan, S.J. Russell, A. Ng, "Distance Metric Learning with Application to Clustering with Side-Information," NIPS, 2002.
- [4] J. Lu, J. Hu, J. Zhou, "Deep Metric Learning for Visual Understanding: An Overview of Recent Advances," IEEE Signal Process. Mag. 2017, 34, 76–84.
- [5] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards real-time object detection with region proposal networks," NIPS, 2015.
- [6] K.Q. Weinberger, J. Blitzer, L. Saul, "Distance Metric Learning for Large Margin Nearest Neighbor Classification," NIPS, 2005. ◀