

製品状態，作業箇所，作業骨格情報に基づく 工程作業認識

大島宏友^{1,a)} 吉井崇哲¹ 蚊戸健浩¹ 前川卓也² 浪岡保男¹

概要：モノづくりの現場では，各種作業や製造装置が行っている加工状況を継続的に把握し，計画に対して遅れが発生している作業の原因説明や改善施策を検討・実施することで生産性の向上を狙っている．しかし，複数の要素作業から構成されている組立工程においては，作業内容の把握を人手に依存した方法で行っており，作業効率や人員確保の点で大きな課題となっている．そこで本研究では，GNN(Graph Neural Network)を用いて組立工程における要素作業を推定する手法を提案する．本手法は，各要素作業における製品状態，作業者の骨格情報，および製品に対する作業者の両手作業位置を統合的に用いて要素作業を推定することが可能である．実際の組立工程で行われた作業を観察したデータを用いて本手法を評価して提案手法の有効性を確認した．

キーワード：組立工程，製品状態，両手作業箇所，骨格情報，GNN

Work Recognition based on Product States, Working Points and Skeleton Data

HIROTOMO OSHIMA^{†1} TAKANORI YOSHII^{†1}
KATO TAKEHIRO^{†1} TAKUYA MAEKAWA^{†2}
YASUO NAMIOKA^{†1}

Abstract: In a manufacturing, to improve productivity, work contents which are done by workers or machines are continuously checked, and some causes of work delay against the plan are traced and analyzed. However, an assembly process consists of more than 10 elemental works, and the data of the work contents are recorded manually. That is an issue in terms of work efficiency and personnel resources. Therefore, we propose a method which recognizes elemental works of an assembly process by using Graph Neural Network. This method enables to estimate them using product states, working points of worker's hands, and worker's skeleton data of each elemental work integrally. Accuracy and validity of this method was investigated with actual data which are collected at an assembly process.

Keywords: Assembly process, Product states, Working points, Skeleton data, GNN

1. はじめに

モノづくりの現場では，AI・IoT 技術の発展に伴い，スマートファクトリー化が進んでいる．作業現場の様々なデータを取得し，データ処理技術を活用して分析した結果を，再び作業現場へフィードバックすることで，生産性向上を狙っている．その中でも，生産能力，歩留まり，作業進捗などは生産性の向上に直結する指標であり，これらを分析・把握して生産管理を推し進めていくことは極めて重要である．

生産管理に必要な情報は様々存在する．機械による作業が中心の作業現場では，設備に取り付けられたセンサから生産管理に必要な情報を直接取得することができる．その一方で，複数の要素作業から構成される組立工程が中心の作業現場では，複雑で変則的な作業に対応するために，人間の作業者に大きく依存している．そのため，生産管理に

必要な情報を取得するには，作業者の作業内容を正確に把握することが必要となる．

現状，作業者の作業内容を把握する方法として，作業者自身が作業情報を記録する方法や，作業監督者を準備して作業情報を記録する方法が行われている．しかし，作業者が作業を一時中断して作業情報を記録することによる作業性の低下や，作業監督者を新たに準備することによる人員コストの点が大きな課題となっている．

近年では，作業者が装着する加速度計などのセンサや，カメラ画像などから得られる作業者の動きを時系列データとして記録し，これらの情報を基にした行動認識技術の研究が盛んに行われている[1]-[7]．本研究においても，これまでに，組立工程における要素作業を認識する手法として，製品の組立状態，および作業者の両手作業箇所の情報から要素作業を推定するしくみを構築してきた[8][9][10]．しか

1 株式会社東芝 生産技術センター.
2 大阪大学大学院情報科学研究科
a) hirotomo1.oshima@toshiba.co.jp

し、IE（インダストリアル・エンジニアリング）の観点から、要素作業をより細かく分割して分析することへのニーズが高く、詳細に分割した要素作業の推定を可能にするしくみの構築が求められている。

本稿では、組立工程における、より詳細な要素作業を認識することを目的とし、従来の製品状態、作業者の両手作業箇所に加えて、新たに作業者の骨格情報を統合的に活用して要素作業を推定する手法について報告する。

2. 対象とする製造現場

近年、消費者ニーズの多様化や製品ライフサイクルの短期化に伴う多品種少量生産が主流となっており、モノづくりの現場では、それに適したセル生産方式が広く採用されている。図1に、東芝の工場の組立工程におけるセル生産のイメージを示す。セル生産では、1人の作業者が所定のエリアで、製品の組立やネジ締め、検査など製品またはユニットを完成させるのに必要な複数の作業を行う。本稿では、これらの作業を要素作業と呼ぶ。ライン生産に比べて、生産機種の変更による段取り時間や、他の作業者の作業遅れによる手待ち時間を低減できる点が特徴である。

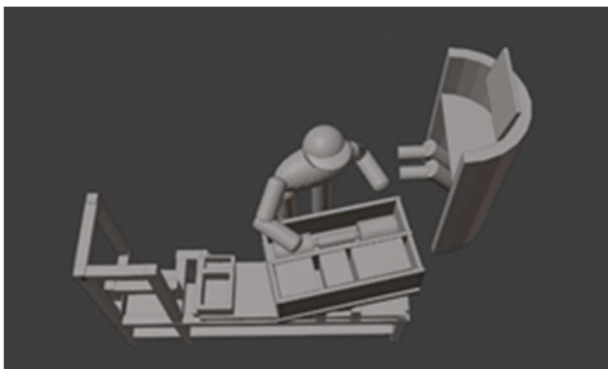


図1 セル生産のイメージ図

セル生産を行っている作業者の作業内容を識別するには課題がある。セル生産は、中型・大型製品を対象とする場合が多く、1つの製品を完成させるのにかかる作業時間が数十分から数時間におよぶなど、長期化する傾向がある。また、作業者の作業時の立ち位置は基本的に一定で、両手を伸ばして部品・工具を取り、適宜姿勢を変えながら作業を行っている。そのため、組立工程の中に類似の姿勢・位置での作業が複数存在し、それらを正確に識別することが困難であることが課題となっている。本研究では、セル生産の組立工程を対象とし、課題の解決手法を検討した。

3. 関連研究

これまでも作業認識に関する様々な手法が提案されている。例えば、ウェアラブルデバイスから取得した作業者のデータを活用して作業認識する手法に関する研究[1]-[4]が報告されている。この研究では、作業者に装着した加速度センサから時系列データを取得し、作業ごとの時系列波形を基に作業認識を行う。しかし、類似した作業が含まれる場合には時系列波形に特徴的な差異が現れにくく、作業内容を正しく識別することが難しい。加えて、例えば同じネジ締め作業であっても、開始直後に行われるネジ締め作業と、終了直前で行われるネジ締め作業を識別することは困難である。

また本研究では、組立工程における製品状態と作業者の両手作業箇所の情報を基に、要素作業を推定する手法を構築している[8][9][10]。この手法では、作業映像から深層学習を用いて推定した製品状態と、オプティカルフローとパターンマッチングの組み合わせにより推定した作業者の両手作業箇所の情報を一定時間観察しながら、どの要素作業であるかをルールベースで推定する。14種の要素作業から構成される実際の組立工程を対象に評価実験を実施し、良好な結果が得られているが、より細かい要素作業の推定を可能にする手法検討や、人間系での調整が必要なルールベースの判定ロジックなど、改良の余地がある。

最近では、人物の骨格情報を高精度かつ自動で抽出できる手法が確立された[11]ことにより、GNN（Graph Neural Network: グラフニューラルネットワーク）を活用した動作認識の研究が注目を集めている[5][6][7]。GNNでは、抽出した骨格情報をグラフ構造として入力し、動作認識を行う。その際、ノードが関節座標、エッジが関節間のつながりに対応しており、各関節間の隣接関係を考慮して学習することができるため、複雑な動作の認識も可能となる。

GNNは、隣接行列と特徴行列から定義される。この2つの行列を掛け合わせることで、隣接しているノードの特徴量を畳み込んだ特徴量を獲得することができる。GNNにおけるConvolution層、および全結合層での処理は、以下の数式(1)、(2)で表すことができる。なお、活性化関数としてReLU関数を用いた場合を示す。

$$v_i^{conv(t)} = \text{ReLU} \left(\sum_{v_j^{conv(t-1)} \in A(i)} W(t) \cdot v_j^{conv(t-1)} \right)$$
$$v_j^{conv(0)} = v_j$$
$$\text{ReLU}(x) = \begin{cases} x & (x \geq 0) \\ 0 & (x < 0) \end{cases} \quad \text{式(1)}$$

$$\text{全結合層} = \sum_{i=1}^N v_i^{conv(t)} \quad \text{式(2)}$$

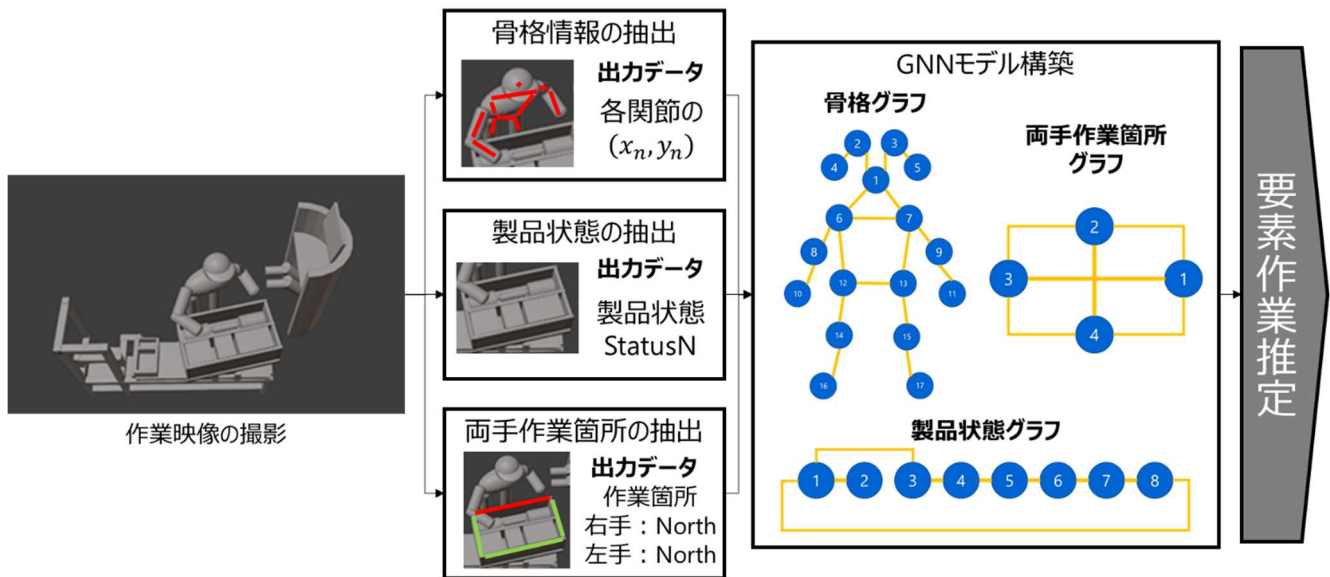


図 2 提案手法の概要

v_i は、 i 番目のノードの初期特徴ベクトル、 $v_i^{conv(t)}$ は、 i 番目のノードを t 回 Convolution 層で畳み込み処理を行った際に更新された特徴ベクトル、 $A(i)$ は隣接行列 A における i 番目のノードおよびそれと隣接しているノードの特徴ベクトルの集合、 $W(t)$ は t 回更新された重みである。

式(1)の処理では、 i 番目のノードおよび隣接しているノードを用いて、 i 番目のノードの特徴ベクトルを更新する。それぞれのノードに重みをかけたものを加算して1つのベクトルとし、ReLU 関数を適用することで新たな特徴ベクトル $v_i^{conv(t)}$ に更新する。畳み込み処理によって t 回更新した N 個の特徴ベクトルは、出力層の直前で式(2)の処理が行われる。式(2)では、更新された N 個の特徴ベクトルを加算して1つのベクトルとし、グラフ構造全体を表す最終的な特徴ベクトルとして出力する。

しかし、GNN を適用した場合でも、類似の作業においては骨格情報に特徴的な差異が現れにくいため、更なる改良が検討されている。

4. 提案手法

本稿では、組立工程における製品状態、作業者の両手作業箇所、作業者の骨格情報に対して GNN を適用し、統合的に要素作業を推定する手法を提案する。作業画像から取得できる情報は、製品状態、作業者の両手作業箇所、作業者の骨格情報があるが、これらの情報単体からでは要素作業の推定は困難である。人間系で要素作業を特定する場合には、製品の組付け状態や作業者の両手の作業箇所、作業者の姿勢・位置などから統合的に判断している。同じネジ締め作業でも、製品の組付け状態の違いや作業者がどこの

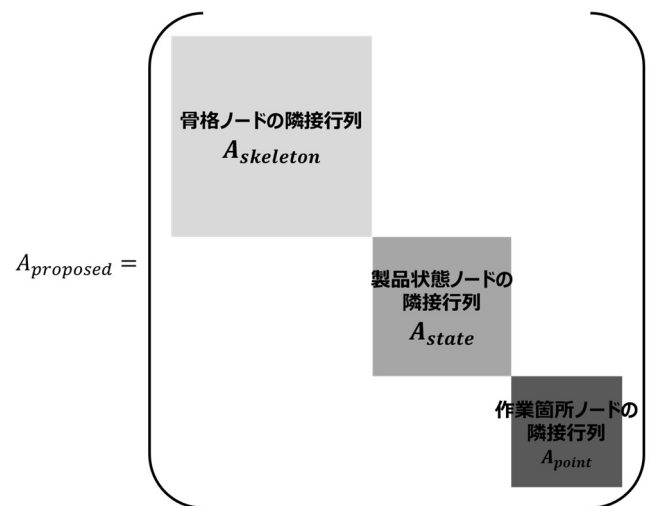


図 3 提案手法の隣接行列

箇所ではネジ締めを行っているかで異なる要素作業に識別している。このことから、作業者の骨格情報に加え、製品状態、作業者の両手作業箇所の3つの情報に対して GNN を適用することを検討した。

図2に提案手法の概要図を示す。撮影した作業映像から、作業者の骨格情報の抽出、製品状態の抽出、および作業者の両手作業箇所の抽出を並行して行い、抽出した値をあらかじめ定義した各グラフ構造に入力することで要素作業を推定する。骨格情報は関連研究に倣ってグラフ構造を定義した。製品状態は、時間の経過とともにその外観が変化することから、状態遷移のグラフとして表現することができ、作業者の両手作業箇所も、作業者が作業する箇所をあらかじめ定義することでグラフとして表現可能である。

また、GNN では、1つの隣接行列内で複数のグラフ構造

を定義することができる。図3に示すように、作業者の骨格グラフの隣接行列 $A_{skeleton}$ 、製品状態グラフの隣接行列 A_{state} 、作業者の両手作業箇所グラフの隣接行列 A_{point} を定義することで、互いに隣接していない3つのグラフを表現することができる。本稿では、この隣接行列を $A_{proposed}$ とする。

以下では、各グラフ構造、特徴行列、および定式化について説明する。

4.1 骨格グラフ

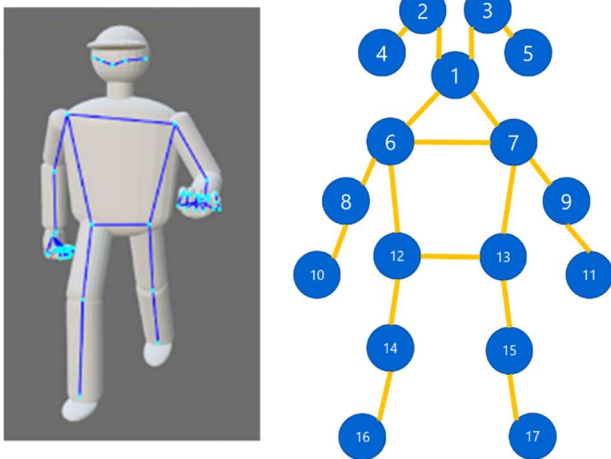


図4 提案手法の骨格グラフ

図4に定義した骨格グラフを示す。骨格グラフは、作業者の関節をノード、関節間のつながりをエッジで表現している。本研究で定義した骨格グラフのノードは左右の目、肩、腕、足など、その数は全部で17個であるが、関節間のつながりが定義可能であれば、任意のノードを用いた骨格グラフに拡張できる。

4.2 製品状態グラフ

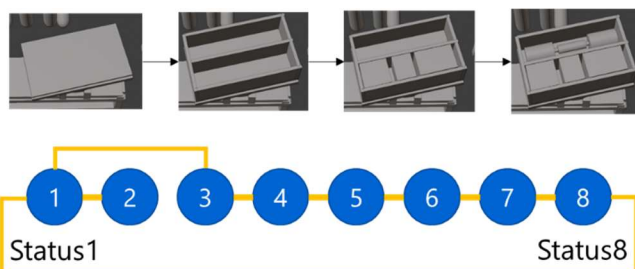


図5 提案手法の製品状態グラフ

図5に、定義した製品状態グラフを示す。製品状態グラフでは、製品状態をノード、製品状態の遷移関係をエッジで表現している。本研究で対象としている組立工程では、

製品状態を Status1 から Status8 までの 8 つに定義することができるためノード数を 8 個としているが、状態遷移の関係が定義可能であれば、任意のノードを用いた製品状態グラフに拡張できる。

4.3 両手作業箇所グラフ

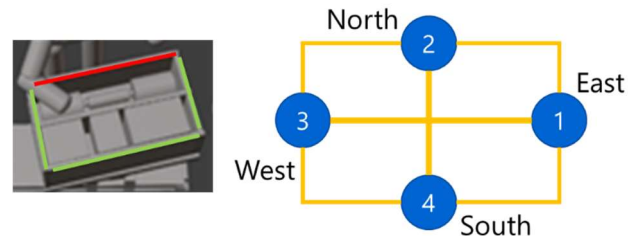


図6 提案手法の両手作業箇所グラフ

図6に、定義した両手作業箇所グラフを示す。両手作業箇所グラフでは、東西南北で定義した作業箇所をノード、隣接する作業箇所をエッジで表現している。本研究で対象としている製品は矩形なので、ノード数を東西南北の4個としているが、隣接関係が定義可能であれば、任意のノードを用いた両手作業箇所グラフに拡張できる。

4.4 特徴行列

骨格グラフ、製品状態グラフ、両手作業箇所グラフのノード特徴量は、5次元ベクトルとした。以下の式(3)に示すように、骨格情報の X, Y 座標、製品状態、右手作業箇所、左手作業箇所構成されている。製品状態は、製品の Status と同じ番号のノードを 1 に、それ以外を 0 にしている。また両手作業箇所も、作業しているノードを 1 に、それ以外のノードを 0 にしている。例えば製品状態が Status3 で、右手作業箇所、左手作業箇所がともに North の場合、骨格グラフの 2 番目のノードの特徴ベクトルは(719.26, 257.81, 0, 0, 0), 製品状態グラフの 3 番目のノードの特徴ベクトルは(0, 0, 1, 0, 0), 両手作業箇所グラフの 2 番目のノードの特徴ベクトルは(0, 0, 0, 1, 1)となる。

$$\text{特徴ベクトル} = \begin{pmatrix} \text{骨格情報の X 座標} \\ \text{骨格情報の Y 座標} \\ \text{製品状態}(0,1) \\ \text{右手作業位置}(0,1) \\ \text{左手作業位置}(0,1) \end{pmatrix} \quad \text{式(3)}$$

4.5 定式化

提案手法における Convolution 層、および全結合層での処理は、以下の数式(4), (5)で表すことができる。なお、活

表 1 要素作業の例

要素作業名	作業内容
要素作業1	吊耳取付け
要素作業2	ネジ止め
要素作業3	ヒートシンク調整
要素作業4	側面部品取付け
要素作業5	ネジ止め
要素作業6	ファン取付け
要素作業7	作業履歴記入
要素作業8	底面部品取付け
要素作業9	ネジ止め
要素作業10	ヒートシンク取付け
要素作業11	ネジ止め
要素作業12	モータ取付け
要素作業13	ファン部品取付け
要素作業14	カバー板取付け
要素作業15	ネジ止め
要素作業16	外観検査

活性化関数として ReLU 関数を用いた場合を示す。

$$v_i^{conv(t)} = \text{ReLU} \left(\sum_{v_j^{conv(t-1)} \in A(i)_{proposed}} W(t) \cdot v_j^{conv(t-1)} \right)$$

$$v_j^{conv(0)} = v_j$$

$$\text{ReLU}(x) = \begin{cases} x(x \geq 0) \\ 0(x < 0) \end{cases} \quad \text{式(4)}$$

$$\text{全結合層} = \sum_{i=1}^N v_i^{conv(t)} \quad \text{式(5)}$$

v_i は、 i 番目のノードの初期特徴ベクトル、 $v_i^{conv(t)}$ は、 i 番目のノードを t 回 Convolution 層で畳み込み処理を行った際に更新された特徴ベクトル、 $A(i)_{proposed}$ は提案手法の隣接行列における i 番目のノードおよびそれと隣接しているノードの特徴ベクトルの集合、 $W(t)$ は t 回更新された重みである。

この式(4)の $v_i^{conv(t)}$ は、図 3 に示した隣接行列の特徴を踏まえると、以下の式(6)と等価である。

$$v_i^{conv(t)} = \begin{cases} \text{ReLU} \left(\sum_{v_j^{conv(t-1)} \in A(i)_{skeleton}} W(t) \cdot v_j^{conv(t-1)} \right) \\ \text{ReLU} \left(\sum_{v_j^{conv(t-1)} \in A(i)_{state}} W(t) \cdot v_j^{conv(t-1)} \right) \\ \text{ReLU} \left(\sum_{v_j^{conv(t-1)} \in A(i)_{point}} W(t) \cdot v_j^{conv(t-1)} \right) \end{cases} \quad \text{式(6)}$$

$A(i)_{skeleton}$ は骨格グラフにおける i 番目のノードおよびそれと隣接しているノードの特徴ベクトルの集合、 $A(i)_{state}$ は製品状態グラフにおける i 番目のノードおよびそれと隣接しているノードの特徴ベクトルの集合、 $A(i)_{point}$ は両手作業箇所グラフにおける i 番目のノードおよびそれと隣接しているノードの特徴ベクトルの集合である。また、 $I_{skeleton}$ は骨格グラフのノード番号の集合、 I_{state} は製品グラフのノード番号の集合、 I_{point} は両手作業箇所グラフのノード番号の集合である。

式(6)の処理では、 i 番目のノードおよび隣接しているノードを用いて、 i 番目のノードの特徴ベクトルを更新する。それぞれのノードに重みをかけたものを加算して1つのベクトルとし、ReLU 関数を適用することで新たな特徴ベクトル $v_i^{conv(t)}$ に更新する。この時、 i の値によって異なる隣接行列を適用する。本研究では、グラフの総ノード数 N は骨格グラフで $I_{skeleton}=17$ 、製品状態グラフで $I_{state}=8$ 、両手作業グラフで $I_{point}=4$ の計 29 個であり、 $i=12$ であれば、 $A_{skeleton}$ を用いて隣接するノードの集合を算出する。そのため、定義した3つのグラフは畳み込み処理が行われる際に、互いに影響を及ぼしあうことはなく、それぞれ独立した畳み込み処理が行われる。畳み込み処理によって t 回更新した N 個の特徴ベクトルは、出力層の直前で式(5)の処理が行われる。式(5)では、更新された N 個の特徴ベクトルを加算して1つのベクトルとし、グラフ構造全体、すなわち、1つの要素作業を表す特徴ベクトルとして出力する。

このことから、作業者の骨格情報、製品状態、作業者の両手作業箇所の特徴量をそれぞれ畳み込んで学習し、それらの結果から統一的に要素作業を推定することが可能となる。

5. 評価

5.1 データセット

空調機の組立工程を対象とし、異なる作業員 9 名分の作業映像に対して要素作業を推定する評価実験を行った。各

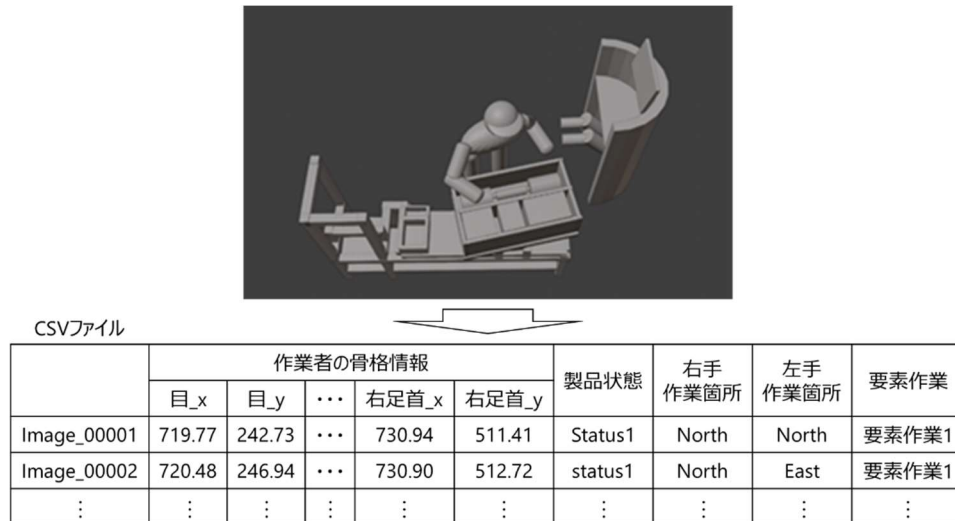


図 8 学習・評価データの作成手順

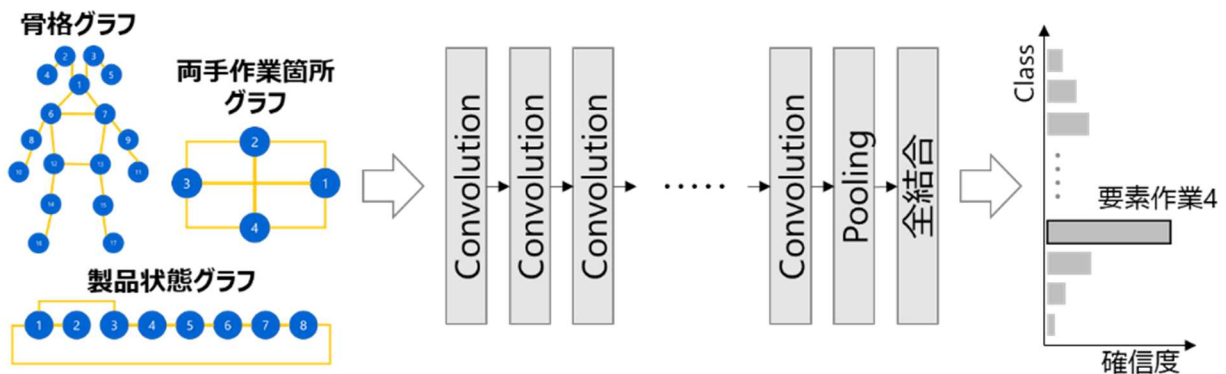


図 7 提案手法の GNN モデル

作業員に対して、2 サイクル分の作業映像を撮影した。この組立工程は、1 サイクルあたり約 15 分で、16 種の要素作業から構成されており、表 1 に要素作業の例を示す。

学習および評価に用いたデータは、図 7 に示すように、作業映像を分割した各画像における製品状態、作業員の両手作業箇所、作業員の骨格情報、および要素作業である。作業員の骨格情報は自動推定ソフトを用いて抽出しているが、抽出結果にはノイズが入ることがある。画面内に複数の作業員が映り込む場合や作業員ではない対象を検出してしまふことが起こりうる。これらのノイズに関しては、データを目視で確認して削除した。また、製品状態、作業員の両手作業箇所、および要素作業については、一枚ずつ画像を目視で確認してラベリングを行った。画像 1 枚に対して、骨格情報のデータが 17 関節の x 座標および y 座標の値で 34 個、製品状態が 1 個、両手作業箇所が 2 個、要素作業が 1 個の合計 38 個のデータがあり、これらのデータを CSV ファイルで準備した。

学習および評価に用いた GNN モデルは、図 8 に示すように、Convolution 層を 15 層、Pooling 層を 1 層組み合わせ

て構成しており、製品状態、作業員の両手作業箇所、作業員の骨格情報を基に、16 クラスの要素作業の中で最も確信度の高い要素作業を推定結果として出力するネットワークとした。学習 Epoch 数は 300 とした。また、撮影した作業映像の 1 サイクル目を学習用、2 サイクル目を評価用とし、画像の枚数はそれぞれ約 20000 枚程度を使用した。評価の指標としては、正解率、Recall (再現率)、Precision (適合率)、F 値を用いた。なお、正解率、Recall、Precision、F 値は以下の式(7)、式(8)、式(9)、式(10)で算出した。学習・評価のプログラムは PyTorch を用いて実装し、GPU 上で計算処理を実施した。学習時の計算処理には、1 Epoch あたり、約 10~15 分ほど要した。

$$\text{正解率} = \frac{\text{正解と GNN 推定が一致している画像枚数}}{\text{全画像枚数}} \quad \text{式(7)}$$

$$\text{Recall} = \frac{1}{16} \sum_{i=1}^{16} \frac{\text{GNN 推定が正解している画像枚数}}{\text{正解の要素作業 } i \text{ の画像枚数}} \quad \text{式(8)}$$

表 3 作業員 A における要素作業の推定結果

		推定																
		要素作業 1	要素作業 2	要素作業 3	要素作業 4	要素作業 5	要素作業 6	要素作業 7	要素作業 8	要素作業 9	要素作業 10	要素作業 11	要素作業 12	要素作業 13	要素作業 14	要素作業 15	要素作業 16	合計
正解	要素作業 1	96	19	3	11	0	0	0	0	0	0	0	0	0	0	0	0	129
	要素作業 2	9	527	29	9	0	0	0	0	0	0	0	0	0	0	0	0	574
	要素作業 3	5	20	1057	0	0	0	0	0	0	0	0	0	0	0	0	0	1082
	要素作業 4	4	1	0	110	0	0	0	0	0	0	0	0	0	0	0	0	115
	要素作業 5	0	0	0	8	441	12	0	0	0	0	0	0	0	0	0	0	461
	要素作業 6	0	0	0	6	16	132	38	4	0	0	0	0	0	0	0	0	196
	要素作業 7	0	0	0	0	1	11	506	25	2	0	0	0	0	0	0	0	545
	要素作業 8	0	0	0	0	0	6	115	305	121	0	0	0	0	0	0	0	547
	要素作業 9	0	0	0	0	0	0	0	246	1135	1	0	0	0	0	0	0	1382
	要素作業 10	0	0	0	0	0	0	0	16	56	760	174	17	0	0	0	0	1023
	要素作業 11	0	0	0	0	0	0	0	0	0	175	1238	16	0	0	0	0	1429
	要素作業 12	0	0	0	0	0	0	0	0	0	9	23	681	440	33	0	3	1189
	要素作業 13	0	0	0	0	0	0	0	0	0	0	0	75	1298	45	0	0	1418
	要素作業 14	0	0	0	0	0	0	0	0	0	0	0	3	9	30	4	137	183
	要素作業 15	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1307	140	1447
	要素作業 16	1	0	5	0	0	0	0	0	0	0	0	0	0	6	270	3088	3370
	合計		115	567	1094	144	458	161	659	596	1314	945	1435	792	1747	114	1581	3368

表 2 作業員ごとの正解率, Recall, Precision, F 値

	作業員 A	作業員 B	作業員 C	作業員 D	作業員 E	作業員 F	作業員 G	作業員 H	作業員 I	平均
正解率	0.8423	0.8143	0.7638	0.8964	0.8800	0.8777	0.8712	0.8857	0.8558	0.8541
Recall	0.7884	0.7219	0.7439	0.8079	0.7373	0.8251	0.7827	0.8111	0.7564	0.7750
Precision	0.7936	0.7348	0.6880	0.8279	0.8378	0.8678	0.8124	0.8584	0.8488	0.8076
F 値	0.7858	0.723	0.7284	0.8158	0.7591	0.8363	0.7912	0.8268	0.7803	0.7830

$$\text{Precision} = \frac{1}{16} \sum_{i=1}^{16} \frac{\text{GNN 推定が正解している画像枚数}}{\text{GNN 推定が要素作業 } i \text{ の画像枚数}} \quad \text{式(9)}$$

$$\text{F 値} = \frac{2 * \text{Recall} * \text{Precision}}{\text{Recall} + \text{Precision}} \quad \text{式(10)}$$

5.2 結果

表 2 に、作業員 A の、正解に対する要素作業推定の結果を示す。縦軸は正解の要素作業、横軸が要素作業推定の結果であり、灰色で示している部分が正解と推定結果が一致している画像枚数を示している。表 2 から、正解の要素作業に対して推定した要素作業が一致している枚数が多く、

高い精度で要素作業を推定できていることがわかる。また、推定結果のバラつきは、前後の要素作業に多い傾向が見てとれる。これは、製品状態を考慮した学習・推定を行えていることを示しており、統合的に推定するモデルが構築できていると考えられる。

作業員 B～作業員 I についても同様に推定結果を集計し、表 3 に作業員ごとの正解率, Recall, Precision, F 値, およびそれらの平均値を示す。全作業員の平均の正解率は 0.8541, 平均の Recall は 0.7750, 平均の Precision は 0.8076, 平均の F 値は 0.7830 であった。製品状態および作業員の両手作業箇所を一定時間観測し、どの要素作業であるかルールベースで推定する従来手法では、作業員 1 人に対して評価実験を行っており、表 4 に示す結果となっている。比較してみると、ほぼ同等の結果が得られていることが分かる。従来手法と提案手法の評価対象は同一の組立工程で

表 4 従来手法と提案手法の比較結果

	正解率	Recall	Precision	F値
従来手法	0.8690	0.7940	0.8260	0.7860
提案手法	0.8541	0.7750	0.8076	0.7830

あるが、要素作業の分割数は14種から16種へ増やしている。これは、現場導入を見据えた担当者とのディスカッションにおいて、要素作業をさらに細かく見直すことへの要望があり、IE分析者、現場担当者とともに要素作業を再定義したためである。このことを踏まえると、要素作業をさらに細かく分割した上で従来手法と同等の推定が可能であることから、要素作業の推定に十分な精度であると考えられる。

要素作業の推定を誤った原因としては、学習データの網羅性が挙げられる。1サイクル目を学習用、2サイクル目を評価用データとして使用しているが、1サイクル目と2サイクル目で標準作業通りの作業を実施していても全く同じ動作を行うことはなく、多少の誤差が生じる。そのため、学習していない製品情報、作業者の両手作業箇所、作業者の骨格情報の組み合わせを評価することになり、誤推定が生じたと考えられる。

また、作業員Cの正解率が他の作業員の正解率に比べて低い原因として、“標準外作業”が発生したことが挙げられる。作業員Cの2サイクル目において、ネジ締め作業中にネジを床に落とす場面があった。このような“標準外作業”は学習用データには含まれていないため、特に誤推定が起りやすい。

推定精度を向上させる施策として、学習データ数を増やすことが効果的であると考えられる。今回使用した学習画像は1サイクル分であったが、複数サイクル分学習させることで、より網羅的になると考えられる。また、今回使用した学習・評価モデルは、convolution層を15層、Pooling層を1層組み合わせで構成しているが、他の学習・評価モデルでの評価は実施できていない。モデルの構造については試行錯誤的な部分があるが、各層の数や順番など改良の余地があると考えている。

6. おわりに

本稿では、複数の要素作業から構成される組立工程において、製品状態、作業員の両手作業箇所、作業員の骨格情報をGNNに入力することで、統合的に要素作業を推定する手法を提案した。1つの隣接行列内で製品状態、作業員の両手作業箇所、作業員の骨格情報のグラフ構造を定義しており、それぞれ独立した畳み込み学習と、それらの結果

から統合的に要素作業を推定することが可能である。作業員9名の作業映像に対して提案手法の評価を行ったところ、平均85.41%の精度で推定できることを確認した。製品状態、作業員の両手作業箇所の情報を一定時間観察し、どの要素作業であるかをルールベースで推定する従来手法と比べると、さらに要素作業を細かく分割した組立工程に対して、同等の精度で推定が可能であり、要素作業の推定に十分な精度であることが確認できた。

今回得られた結果から、提案手法は現場作業の分析に十分活用可能である。製造現場への導入を見据えた今後の展望として、要素作業の推定精度向上に向けたモデルの改良、および標準外作業を自動で検出する手法の検討を進めていき、提案手法の更なる発展を狙っていく。

参考文献

- [1] Morales, J., Yoshimura, N., Xia, Q., Wada, A., Namioka, Y. and Maekawa, T. Acceleration-based Human Activity Recognition of Packaging Tasks Using Motif-guided Attention Networks, Proc. of IEEE Int'l Conf. on Pervasive Computing and Communications (PerCom 2022)
- [2] Xia Qingxin, et al. “Robust Unsupervised Factory Activity Recognition with Body-worn Accelerometer Using Temporal Structure of Multiple Sensor Data Motifs,” in Wearable Ubiquitous Technol, 2020.
- [3] Hammerla, N. Y., Halloran, S. and Ploetz, T.: Deep, Convolutional, and Recurrent Models for Human Activity Recognition using Wearables, CoRR, Vol. abs/1604.08880 (2016).
- [4] 吉村 直也, 前川 卓也, 原 隆浩, 和田 篤, 浪岡 保男. 工場・物流センタにおける作業順序を考慮した少量学習データでの作業行動認識手法の検討. IPSJ SIGUBI 情報処理学会 ユビキタスコンピューティングシステム研究会 第74回(2022)
- [5] Y. Sijie, X. Yuanjun, and L. Dahua, “Spatial Temporal Graph Convolutional Networks for Skeleton Based Action Recognition,” in Association for the Advancement of Artificial Intelligence, 2018.
- [6] C. Li, Z. Cui, W. Zheng, C. Xu, R. Ji, and J. Yang, “Action-Attending Graphic Neural Network,” in Institute of Electrical and Electronics Engineers, pp.3657-3670, 2018.
- [7] Ibsa K. Jalata, Thanh-Dat Truong, Jessica L. Allen, Han-Seok Seo, Khoa Luu, Movement Analysis for Neurological and Musculoskeletal Disorders Using Graph Convolutional Neural Network, MDPI, Future Internet 2021
- [8] 吉井 崇哲, 大島 宏友, 白川 悠太, 西村 圭介, 蚊戸 健浩, 前川 卓也, 浪岡 保男. 製品の状態推定と作業箇所推定に基づく工程作業認識. 情報処理学会 第84回全国大会(2022)
- [9] 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. 機械学会生産システム部門研究発表講演会 2022 講演論文集
- [10] 白川 悠太, 大島 宏友, 吉井 崇哲, 西村 圭介, 蚊戸 健浩, 前川 卓也, 浪岡 保男, 人物姿勢推定および密なオブティカルフローに基づく作業箇所推定. 情報処理学会 第84回全国大会(2022)
- [11] C. Zhe, S. Tomas, W. Shih-En, and S. Yaser, “Real-time Multi-person 2D Pose Estimation Using Part Affinity Fields,” in Computer Vision and Pattern Recognition, pp.1302-1310, 2017.