作業者骨格・製品状態・作用点の情報からの作業認識

Work recognition using information of operator skeleton, product states, and processing parts

浪岡 保男 [1*] 大島 宏友 [2] 新沼 歩 [2] 福田 雅允 [2] 前川 卓也 [3]

Yasuo Namioka[1*] Hirotomo Oshima[2] Ayumu Niinuma[2] Masamitsu Fukuda[2] Takuya Maekawa[3]

[1] 東京都立産業技術大学院大学 Graduate School of Industrial Technology, Advanced Institute of Industrial Technology
[2] 株式会社東芝 生産技術センター Corporate Manufacturing Engineering Center, Toshiba Corporation
[3] 大阪大学大学院 情報科学研究科 Graduate School of Information Science and Technology, Osaka University
*Corresponding author: Yasuo Namioka, namioka-yasuo@aiit.ac.jp

Abstract In a manufacturing, because understanding of work progress is crucial in productivity improvement, efforts are being made to automatically collect work-related sensor data. However, in an assembly process which is mainly done by workers, it is difficult to estimate work activities automatically because of the complexity and variety of them. We thus propose a new method which estimates elemental works of assembly process automatically by integrating information about workers, a product of interest, and processing parts. We apply GNN for the modeling and the effectiveness of this method was investigated with actual data which are collected at an assembly process.

Keywords work recognition; product states; processing parts; skeleton data; GNN

1 はじめに

スマート・マニュファクチャリングに向けた取り組みが，モノづくりの現場で進められている[1]．AI・IoT 技術を活用して各種作業に関わる様々なデータを自動収集・分析し，改善活動に繋げている．こうした中で作業分析も行われており，製造の生産性を向上させる上で重要な取り組みの一つである．作業分析のために収集されるデータは作業進捗の把握においても活用される．生産性に直結する指標であることから，あらゆる製造現場で徹底した管理が行われている[2, 3]．

機械による作業が中心の製造現場では，設備に取り付けたセンサから作業進捗の把握に必要なデータを直接取得することができる[4-6]．一方で，組立作業が中心の製造現場では，その作業内容が複雑かつ多様であることから，作業の自動化が進められている現在でも作業者に大きく依存している．そのため作業進捗を把握するには，作業者の作業内容を正確に捉えることが必要となる．現状，作業者本人や作業監督者が進捗情報を登録する方法が行われているが，作業性や人員コストの点で課題となっている．

近年では，カメラ画像から得られる作業者の骨格の時系列データを記録し，グラフとして表現される骨格から GNN（Graph Neural Network）[7,8]を用いて行動認識を行う研究が進められている[9,10]．しかし，セル生産の組立工程では，同じ姿勢・動作で行う作業が複数存在することから，作業者の骨格のみから作業内容を認識することが難しく，作業対象となる製品の状態や，その製品との作用点の情報を統合的に扱うことが求められる．

そこで本稿では，組立工程における複雑で多様な作業内容を詳細かつ高精度に自動推定することを目的に，作業者の骨格だけでなく，作業対象（製品）の状態や作業者と製品の作用点もグラフ構造で記述し，GNN を用いて要素作業レベルで作業内容を自動推定する新たな手法を提案する．このとき，骨格，製品の状態，作用点の情報を個別にグラフ化してグラフ畳み込みを行っても，それぞれのグラフが保持する情報が交換されず，それらの情報を統合的に用いた作業認識が行えない．そこで本研究では，作業者が製品のどの箇所の作業を手で行っているかの作用点を表すグラフをハブとして，作業者骨格の手のノードと製品状態を表すグラフを連結することで，作業者が特定の状態における製品の特定箇所に対して行っている作業の情報を効果的に抽出する．

本研究の特徴は以下のとおりである．

- 複数の作業情報を統合的に扱うモデルの提案：複数の作業情報を統合的にモデル化し，要素作業を推定する手法を提案する．統合的なモデル化に GNN を用いることで，要素作業の内容を特徴づける各作業情報間の関連性を考慮した学習が可能である．
- セル生産の組立作業現場における実作業データを用いた提案手法の実証：セル生産で製造されている製品の組立工程を対象に，異なる 9 名の作業者の実作業データを用いた評価を実施し，提案手法の有効性を示した．

2 対象とする製造現場

製造現場では，微細なものから社会インフラに関わるものまで，様々な製品を製造している．近年は，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[11]，モノづくりの現場では，それに適したセル生産が広く採用されている．図 1 に，セル生産による，一般的な製品の組立工程のイメージを示す．セル生産では，1 人の作業者が所定のエリアで，部品の取り付け，ネジ締め，外観検査など，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，生産機種の切り替え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[12]．

セル生産を行っている作業者の作業内容を自動で把握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 10 分から数時間におよぶなど，作業時間が長時間化する傾向がある．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．本研究では，セル

Received: 2024-10-30; Published online: 2025-02-28

生産で製造されている様々な製品の組立工程を対象に，課題の解決手法を検討した．

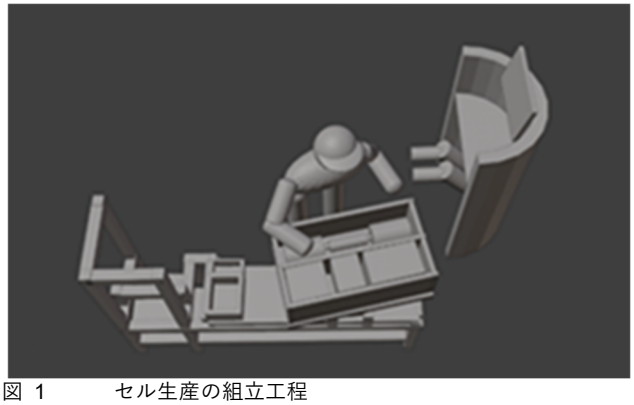


図 1　セル生産の組立工程

表 1　先行研究と本研究の位置づけ

	認識に利用する情報			認識/統合手法
	骨格	製品状態	作業箇所	
文献 [13-24]	○			DTW CNN LSTM
文献 [25-27],[29]		○		オプティカルフロー CNN
文献 [28, 30]		○	○	ルールベース
文献 [9, 10],[31-33]	○			GNN
本研究	○	○	○	GNN

3　関連研究

これまでにも行動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や骨格抽出技術から取得したデータを活用して行動認識する手法に関する研究[13-24]が報告されている．この研究では，人間に装着した加速度センサから時系列データを取得し，動作ごとの時系列波形を基に行動認識を行う．しかし，類似した動作が含まれる場合は時系列波形に特徴的な差異が現れにくく，動作内容を正しく識別することが難しい．

また，作業現場を撮影した動画像から作業の進捗情報を抽出する研究も行われている[25-27]．本研究の先行研究でも，撮影した作業映像に対して，深層学習により推定した製品の組付け状態の情報と，作業者の両手作業位置の情報を基に作業進捗を自動抽出する手法を構築している[28-30]．この手法では，製品状態と両手作業箇所の情報の遷移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．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，作業分割の粒度の点やユーザによる事前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，GNN を活用した行動認識の研究が注目を集めている[31-33]．GNN は，隣接行列と特徴行列から定義され，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 を活用した行動認識では，抽出した人間の骨格位置の情報を特徴量として事前に定義したグラフ構造に入力する．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が可能なため，複雑な動作認識も可能となる．

表 1 に先行研究と本研究の位置づけを示す．これまでに，作業者の骨格情報と対象製品に関わる各種情報を統合的に扱い作業推定を行っている研究は報告されていない．本研究では，先行研究を踏まえ，作業者の骨格位置情報に加えて，製品の組付け状態，作業者の両手作業箇所の情報を，GNN を用いて統合的にモデル化し，学習・推定を行う手法を提案する．

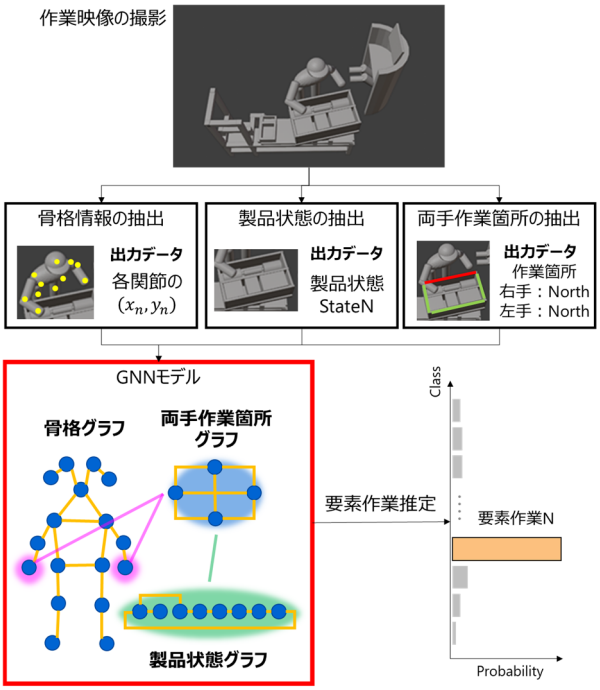


図 2　提案手法の概要

4　提案手法

本稿では，作業者や作業対象，作業者と作業対象の作用点の情報を統合的にモデル化し，要素作業を推定する新たな手法を提案する．具体的には，作業者の骨格位置の情報，製品の組付け状態の情報，作業者の両手作業箇所の情報を用いた．また統合的にモデル化する方法として，GNN を適用した．本章では，提案手法の概要，定義したグラフ構造，特徴行列，学習プロセス，推定プロセスについて説明する．

概要

図 2 に，提案手法の概要を示す．作業現場の上方より撮影した作業映像を静止画に分割し，その静止画 1 枚 1 枚から作業者

の骨格情報，製品状態，および作業者の両手作業箇所の抽出を並行して行う．抽出した値をあらかじめ定義したグラフ構造に入力し，それらの特徴を GNN で学習することで要素作業を推定する．図 2 の赤枠部分が，新たに提案する 3 つの作業情報を統合的に扱う GNN モデルである．

作業内容を識別する情報として，作業者，作業対象，作業者と作業対象の作用点などが挙げられるが，これらの情報単体からでは要素作業の推定が難しい場合がある．人間が作業内容を認識する場合，製品の組付け状態や作業箇所，作業者の姿勢・位置などから統合的に判断している．例えば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箇所でネジ締めを行っているかで作業内容を識別している．また，作業対象によって，作業内容の識別に必要な情報の組み合わせが異なる．作業時間が長く，作業内容が複雑になるにつれて，必要な作業情報の数が増える．これらを踏まえ，組立作業の作業内容の推定に必要な作業情報を統合的にモデル化する本手法を構築した．作業情報をグラフで表現して GNN で処理することで，これらの情報を統合的に考慮した認識を行うことができる．

グラフ構造

GNN で事前に定義するグラフとして，骨格グラフ，製品状態グラフ，両手作業箇所グラフの 3 つのグラフを定義した．また，これらの作業情報間の関係性を考慮し，エッジで 3 つのグラフを結合させた．以降で，それぞれのグラフ構造，3 つのグラフを結合させた提案手法のグラフ構造について説明する．

骨格グラフについて説明する．図 3 は，先行研究[9]に倣って定義した骨格グラフである．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．

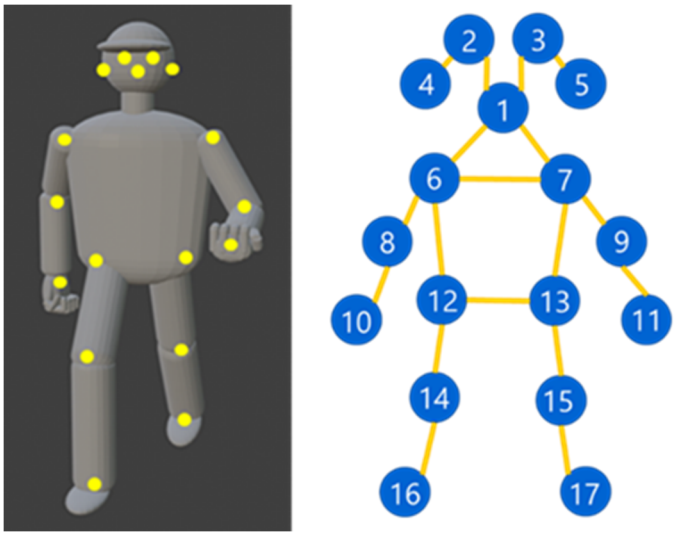


図 3　　骨格グラフ

製品状態グラフについて説明する．図 4 に，製品状態グラフを示す．製品状態は，組立作業とともに変化していく製品の外観に対応し，外観遷移が存在する製品であれば製品状態グラフを定義することが可能である．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 State1 から State8 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．組立作業に伴う製品の外観の遷移が N 段階ある組立工程であれば，N 個のノードを用いて製品状態グラフを定義することができる．

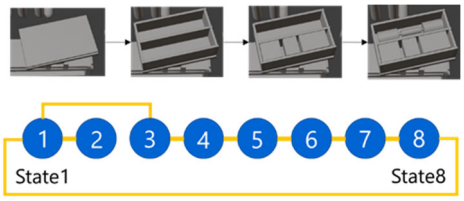


図 4　　製品状態グラフ

両手作業箇所グラフについて説明する．図 5 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．

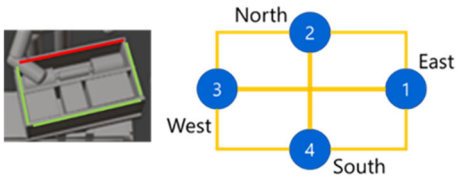


図 5　　両手作業箇所グラフ

統合的に扱う GNN モデルについて説明する．骨格グラフ，製品状態グラフ，両手作業箇所グラフの 3 つを結合させた 1 つのグラフ構造を定義している．作業者は作業対象に対して組付けやネジ締め等の作業を行うことを踏まえると，作業者と作業対象の関係性をグラフ構造内に表現することで，さらに詳細な特徴を学習することができる．そこで，図 6 に示すように骨格グラフの両手のノードと両手作業箇所グラフの全ノード，および，製品状態グラフの全ノードと両手作業箇所グラフの全ノードをエッジで結合させた 1 つのグラフを新たに定義した．骨格グラフと両手作業箇所グラフでは，特に両手の位置と作業者の製品に対する作業箇所に強い関連性があることから，両手のノードと両手作業箇所の全ノードを結合した．また，製品状態グラフと両手作業箇所グラフにおいては，どの製品状態の時にどの箇所で作業をしているかを特徴として捉えるために，お互いの全てのノード同士を結合している．これによって，どのような製品状態の時に，作業者はどのような位置・姿勢で，製品の

どの箇所で作業をしているか，といった極めて細かい特徴を学習することが可能となり，要素作業の推定をより高精度に行うことが可能になる．

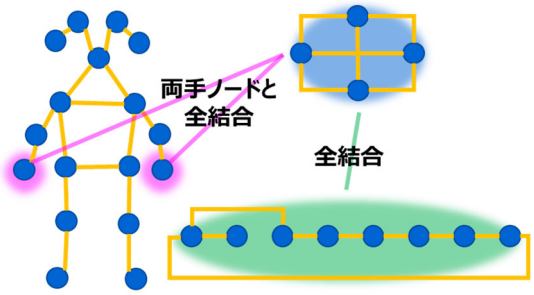


図 6　　統合的に扱うグラフ構造

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，以下の式（1）で示す 5 次元の特徴ベクトルで表現した．特徴ベクトルは，骨格情報の X，Y 座標，製品状態，右手作業箇所，左手作業箇所で構成されている．

骨格情報の抽出では，自動抽出ソフトを活用して各関節の X, Y 座標を取得し，特徴ベクトルに入力している．製品状態，および，両手作業箇所の抽出は目視でラベリングした結果を用いており，製品の State に対応する番号のノードを 1 に，それ以外を 0 にし，両手作業箇所も，作業している箇所に対応するノードを 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 State3 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．

$$\text{特徴ベクトル} = \begin{pmatrix} \text{骨格情報の X 座標} \\ \text{骨格情報の Y 座標} \\ \text{製品状態} \\ \text{右手作業位置} \\ \text{左手作業位置} \end{pmatrix} \tag{1}$$

学習プロセス

本研究では，GNN の中でも ConvGNN（Convolutional Graph Convolutional Network）を採用している．GNN には，ConvGNN（Convolutional Graph Convolutional Network），RecGNN（Recurrent Graph Neural Network），GAE（Graph Autoencoder）などがあるが，ConvGNN は結合しているノード間のみで畳み込み学習を行い，特徴を抽出することができる．3 つの作業情報の結合関係の特徴を詳細に捉えることができると考え，ConvGNN を採用した[7].

各グラフは，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新たな特徴量を獲得することができる．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 (2), (3) 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[8].

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

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

$$\mathrm{ReLU}(x) = \begin{cases} x(x \geq 0) \\ 0(x < 0) \end{cases} \tag{2}$$

$$\mathrm{Output} = \mathrm{Linear}\left(\sum_{i=1}^{N} v_i^{conv(T)}\right) \tag{3}$$

v_iは，i 番目（$i = 1,2,\cdots,N$）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 層目（$t = 1,2,\cdots,T$）の 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 は重みである．

式（2）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 個の特徴ベクトルに対して式（3）の処理を行った結果がモデルの出力となる．式（3）の Linear 関数では，更新された N 個の特徴ベクトルを加算して 1 つのベクトルを生成する．このベクトルの次元数は要素作業のクラス数 M であり，各要素はそれぞれのクラスの確からしさを表す確率値である．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合的に要素作業を推定することが可能となる．

また，提案手法では，図 7 に示すように， Convolution 層を 15 層，Pooling 層を 1 層組み合わせたモデルを用いた．

推定プロセス

推定は，学習した GNN モデルを用いて行う．入力情報は，作業者の骨格情報の X,Y 座標，製品状態，右手作業箇所，左手作業箇所で，式 (3) と同様の特徴ベクトルとして各ノードに与える．また，学習プロセスと同様に，Convolution 層で特徴量の畳み込みを行い，全結合層で要素作業のクラス数 M の次元数を有したベクトルを生成する．このベクトルの中で，最も成分値の高いクラスを要素作業の推定結果として出力する．なお，本稿では，要素作業のクラス数 M は 16 である．

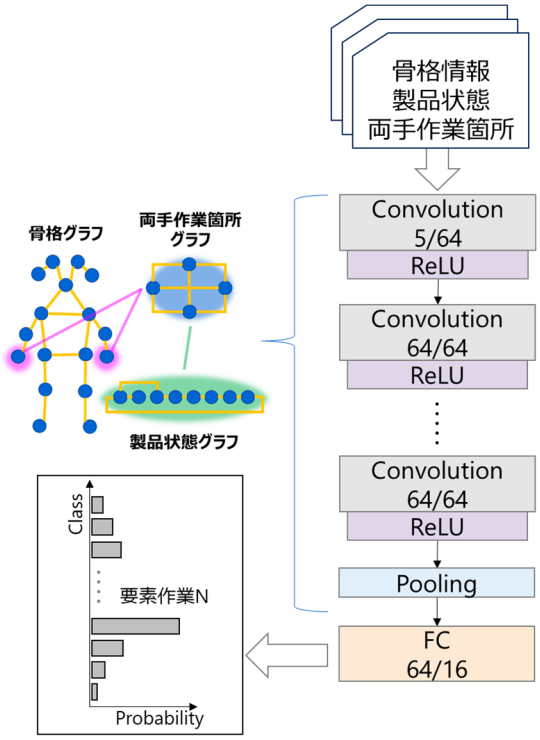


図 7　統合的に扱う GNN の構成

5　評価

実在する一般的なセル生産の組立工程に本提案手法を適用して，有効性を確認した．本章ではその評価内容について述べる．

対象工程

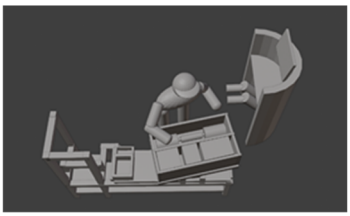
本稿では，異なる作業者 9 名分の作業映像に対して，要素作業を推定する評価実験を行った．この組立工程では産業用の空調機を製造しており，1 サイクルあたり約 15 分，16 種類の要素作業（PM-W1～PM-W16）から構成されている．表 2 に各要素作業の作業内容を示す．

学習・評価データセット

図 8 に学習・評価に使用したデータの構成を示す．学習および評価に用いたデータは，作業映像を分割した各画像における製品状態，作業者の両手作業箇所，作業者の骨格情報，および要素作業である．学習データに関して，作業者の骨格情報は自動推定ソフトを用いて抽出し，製品状態，作業者の両手作業箇所，および要素作業は，一枚ずつ画像を目視で確認してラベリングを行った．評価データに関しても同様に，作業者の骨格情報は自動推定ソフトを用いて抽出し，製品状態，作業者の両手作業箇所は，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．作業映像は各作業者 2 サイクル分撮影しており，1 サイクル目を学習，2 サイクル目を評価用とした．なお，対象としているセル生産の組立工程では，作業者ごとに行う組立工程が異なる．そのため，同一作業者のデータで学習・推定を行った場合の精度を評価することが重要であり，Cross-validation は行っていない．

表 2　要素作業の作業内容

要素作業名	作業内容
PM-W1	吊耳取付け
PM-W2	ネジ止め
PM-W3	ヒートシンク調整
PM-W4	側面部品取付け
PM-W5	ネジ止め
PM-W6	ファン取付け
PM-W7	作業履歴記入
PM-W8	底面部品取付け
PM-W9	ネジ止め
PM-W10	ヒートシンク取付け
PM-W11	ネジ止め
PM-W12	モータ取付け
PM-W13	ファン部品取付け
PM-W14	カバー取付け
PM-W15	ネジ止め
PM-W16	外観検査



CSVファイル

	作業者の骨格情報					製品状態	右手作業箇所	左手作業箇所	要素作業
	目_x	目_y	･･･	右足首_x	右足首_y				
Image_00001	719.77	242.73	･･･	730.94	511.41	Status1	North	North	要素作業1
Image_00002	720.48	246.94	･･･	730.90	512.72	Status1	North	East	要素作業1
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮

図 8　学習・評価データの構成

評価方法

2 通りの評価実験について述べる．

実験 1 では，使用する作業情報の数と推定精度の関係を評価するため，提案手法を用いて以下の 4 つのモデルを作成し推定精度比較を行った．

- モデル 1.　作業者の骨格情報のみを GNN に適用し，要素作業を推定
- モデル 2.　作業者の両手作業箇所，作業者の骨格情報を GNN に適用し，要素作業を推定
- モデル 3.　製品状態，作業者の骨格情報を GNN に適用し，要素作業を推定
- モデル 4.　製品状態，作業者の両手作業箇所，作業者の骨格情報を GNN に適用し，要素作業を推定

実験 2 では，骨格グラフ，製品状態グラフ，両手作業箇所グラフを独立して定義し学習したモデル（モデル 4）と，結合して 1 つのグラフ構造として定義し学習したモデルの推定精度比較を行った．

- モデル 5.　骨格グラフ，製品状態グラフ，両手作業箇所グラフを結合して 1 つのグラフとして定義し，要素作業を推定

各モデルの推定精度指標には F 値を用いた．また，実験 1 では，各モデルの推定結果の分布についても評価を行っており，学習・推定に用いる情報を変化させることで推定精度および推定結果の分布に与える影響を分析した．実装には PyTorch を用いており，GPU 上で計算処理を実施した．

図 9　学習・評価データの構成作業者 9 名における各モデルの F 値比較

結果・考察

実験 1 について，作業者 9 人（Worker A~Worker I）の作業映像に対する，モデル 1，モデル 2，モデル 3，モデル 4 の F 値の比較を図 9 に示す．F 値の平均は，モデル 1 が 0.3869，モデル 2 が 0.3844，モデル 3 が 0.7948，モデル 4 が 0.7830 であった．モデル 1 とモデル 2 の F 値の平均値は同等，モデル 3 とモデル 4 の F 値の平均値も同等であり，モデル 1 とモデル 2 に対してモデル 3 とモデル 4 の F 値の平均値が大きく上回る結果となった．このことから，要素作業の推定において，使用する情報を増やすことで推定精度が向上する傾向にあること，および本稿で対象にした製品に対しては製品状態の情報が推定精度の向上に大きく寄与していることが確認できる．製品状態の推定情報を加えることで，作業の進行具合を概ね特定することができ，作業者の骨格情報，両手作業箇所の情報から，その製品状態における要素作業を詳細に推定することが可能になったと考えられる．対象にする製品によっては作業者の両手作業箇所の情報が極めて重要になることも考えられる．本稿で対象にした製品は両手作業箇所が 4 か所だが，数百から数万か所で作業する製品も存在する．作業箇所の数が膨大になるにつれて，作業箇所の情報から候補となる要素作業の絞り込みが可能となり，高精度な要素作業推定に大きく寄与すると考えられる．

また，各モデルの推定における傾向を確認するために，モデル 1，モデル 2，モデル 3，モデル 4 の要素作業推定結果の混同行列を図 10 に示す．X 軸は推定した要素作業，Y 軸は正解の要

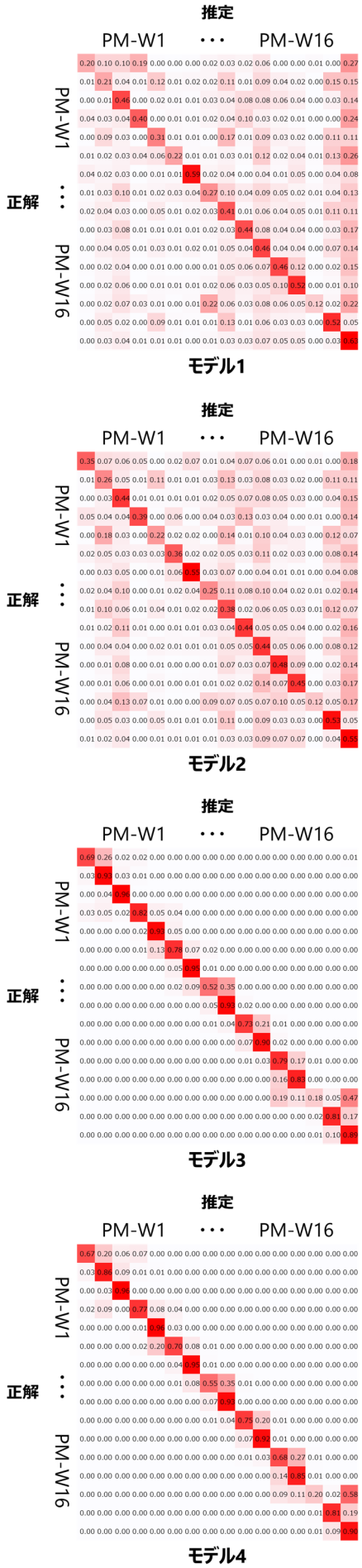


図 10　各モデルの推定結果の混同行列

素作業を表している．この結果からも，製品状態の情報が加わることで誤推定のバラつきが小さくなることが確認できる．製品状態の情報のみでは，同じ製品状態時に行われる作業内容を区別することができず，また，作業者の骨格情報と両手作業箇所の情報のみでは，類似している姿勢・位置で行う他の要素作業と識別することができない．3 つの情報を統合的に活用することで，候補となる要素作業の絞り込みが可能であることを示している．

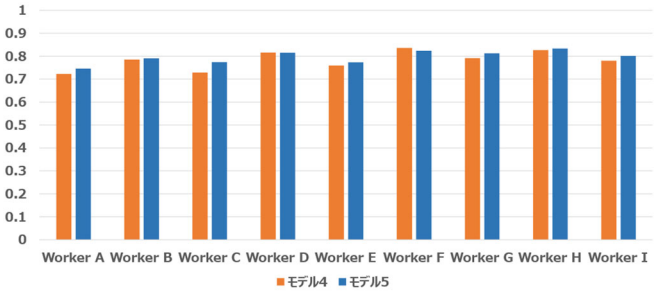


図 11　3 つのグラフ間の結合有無の違いによる F 値比較

実験 2 について，作業者 9 人の作業映像に対する，モデル 4，モデル 5 の F 値の比較を図 11 に示す．モデル 5 の F 値の平均は 0.7966 であり，3 つのグラフを独立して定義したモデル 4 より推定精度が高いことを確認した．セル生産の組立工程を構成する要素作業を正確に推定するには，作業者の情報だけではなく，作業を行っている製造対象や作業者と製造対象の作用点の情報を統合的に学習する必要がある．ある要素作業を行うときに，どのような姿勢で，どのような状態の製品に対して，どのような箇所で作業しているかという作業情報間の関連性をより詳細に捉えるため，提案手法では，骨格グラフの両手ノードと両手作業箇所の全ノード，および製品状態グラフの全ノードと両手作業箇所グラフの全ノードを新たに結合させた．これによって，GNN の畳み込み学習をする際に，作業者の骨格位置，製品状態，作業者の両手作業箇所の情報の 3 つが互いに関連する特徴をより詳細に捉えることが可能になり，推定精度が向上したと考えられる．なお，骨格グラフと製品状態グラフなど，2 つのグラフを結合させて 1 つのグラフとすることも構造上可能だが，結合するノードの選定やグラフ構造全体における説明性の観点から実施していない．

提案手法が十分な推定結果であることを確認したが，さらに推定精度を上げる施策としては，モデル・パラメータの調整や，学習データの増量が効果的と考えられる．

6　まとめ

本稿では，作業者の骨格情報，製品状態，両手作業箇所の情報を統合的にモデル化し，要素作業の推定を行う新たな手法について述べた．統合的なモデル化には GNN を用いており，骨格グラフ，製品状態グラフ，両手作業箇所グラフの 3 つのグラフを結合した 1 つのグラフを定義した．作業者 9 名の作業映像に対して提案手法を評価したところ，最も高い F 値の平均が 0.7966 という結果を得て，3 つの情報を統合的に学習できていることを確認した．また，本稿の対象工程で実際に作業進捗管理を行う担当者の精度を確認すると，1 つの作業の開始時および終了時に平均 3~4 秒ずつの誤差があった．一方で，本手法を適用した場合，推定精度が 0.7966 であったことから，誤推定の割合は約 2 割となる．対象とした組立工程の総作業時間が約 15 分，要素作業数が 16 種類であることから，要素作業 1 つあたりに含まれる誤推定は平均約 11 秒であり，誤推定が作業の開始時と終了時に集中していることを踏まえると，各要素作業の開始時および終了時に平均 5~6 秒の誤差が生じることになる．この誤差は作業進捗管理において許容可能な範囲であり，本手法で代替した場合にも運用可能であることから，十分な推定精度であるといえる．本手法は，作業者が 1 人で組立作業を行うライン生産およびセル生産の作業現場で，外観が組立作業とともに遷移していく製品の工程であれば，汎用的に適用することが可能である．

本手法では，画像 1 枚から作業内容を推定しており，比較的少ない計算量で高速に推定することが可能である．一方で，作業内容を詳細に推定するには，作業の時系列を考慮した推定が効果的であると考えており，今後は本研究をベースに時系列性を考慮したモデルへの改良を進めていく．また，本稿では，製品状態の情報，および両手作業箇所の情報に正解データを用いて評価を行った．今後は，先行研究の手法で推定した値を用いて評価を行う．

参考文献

1. 独立行政法人情報処理推進機構: DX 白書 2023
2. 経済産業省，厚生労働省，文部科学省: 2022 年版　モノづくり白書
3. 経済産業省，厚生労働省，文部科学省: 2023 年版　モノづくり白書
4. Maeda, M., Sakurai, Y., Tamaki, T.: Method for recognizing operation statuses of metal cutting machine tools using spindle motor current, Transactions of the JSME, Vol.84, No.860, 2018(in Japanese)
5. 高村藤寿，西澤泉: 建設機械稼働データを活用した稼働状況の見える化と顧客価値の創造，システム制御情報論文誌, Vol.62, No.8, pp.323-327, 2018
6. 緒方淳，村川正宏，飯田誠: 風力発電スマートメンテナンスのための振動データ解析に基づく状態監視システムの構築，風力エネルギー利用シンポジウム, 37 巻, p.385-388, 2015
7. Zonghan, W., Shirui, P., Fengwen, C., Guodong, L., Chengpi, Z., Philip S., Y.: A Comprehensive Survey on Graph Neural Networks, IEEE Transactions on Neural Networks and Learning Systems, Vol.32, Issue:1, 2021
8. Kipf, T. N., Welling, M.: Semi-Supervised Classification with Graph Convolutional Networks, International Conference on Learning Representations, 2017
9. Obinata, Y., Yamamoto, T.: Temporal Extension Module for Skeleton-Based Action Recognition, 25th International Conference on Pattern Recognition, 2021
10. Komura, K., Horikawa, M., Okamoto, A.: Multimodal Learning of Sensing Data and Skeletal Data for Estimation of Worker Behavior, JIMA, Vol.74, p.31-39, 2023(in Japanese)
11. 総務省: 令和 2 年度版　情報通信白書
12. 小沢浩: セル生産による生産性増大の原理，組織化学論文誌, Vol.38, No.3, p.55-65, 2005
13. 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, Proceedings of the 2022 IEEE International Conference on Pervasive Computing and Communications, pp.1-12(2022).

14. Xia, Q., Korpela, J., Namioka, Y. and Maekawa, T.: Robust Unsupervised Factory Activity Recognition with Body-worn Accelerometer Using Temporal Structure of Multiple Sensor Data Motifs, Proc. ACM Interact. Mob. Wearable Ubiquitous Technol., Vol. 4, No.3(2020).
15. Yoshimura, N., Maekawa, T., Hara, T., Wada, A. and Namioka, Y.: Acceleration-Based Activity Recognition of Repetitive Works with Lightweight Ordered-Work Segmentation Network, Proc. ACM Interact. Mob. Wearable Ubiquitous Technol., Vol. 6, No. 2 (2022).
16. Hammerla, N. Y., Halloran, S. and Ploetz, T.: Deep, Convolutional, and Recurrent Models for Human Activity Recognition using Wearables, Proceedings of the Twenty-Fifth International Joint Conference on Artificial Intelligence, pp.1533-1540 (2016).
17. Zhang, Y., Zhang, Z., Zhang, Y., Bao, J., Zhang, Y. and Deng, H.: Human activity recognition based on motion sensor using u-net, IEEE Access, Vol. 7, pp.75213–75226 (2019).
18. Garcia-Ceja, E., Galv´an-Tejada, C. E. and Brena, R.: Multi-view stacking for activity recognition with sound and accelerometer data, Information Fusion, Vol. 40, pp.45–56 (2018).
19. Cao, Z., Simon, T., Wei, S. and Sheikh, Y.: Realtime Multi-person 2D Pose Estimation Using Part Affinity Fields, IEEE Transaction on Pattern Analysis and Machine Intelligence, Vol.43, pp.172-186(2021).
20. Liu, J., Shahroudy, A., Xu, D., Wang, G.: Spatio-Temporal LSTM with Trust Gates for 3D Human Action Recognition, Computer Vision and Pattern Recognition, 2016
21. Du, Y., Fu, Y., Wang, L.: Skeleton based action recognition with convolutional neural network, 3rd IAPR Asian Conference on Pattern Recognition, 2015
22. Ichihara, K., Takeuchi, M., Kanai, K., Katto, J.: Accuracy Evaluation of Human Motion Recognition Based on Pose Estimation, ITE Winter Annual Convention 2017(in Japanese)
23. Zheng, H., Kato, C., Komatsu, S., Aoki, K.: Motion Analysis of Traditional Korean Dance using Openpose and DTW Distance - Visualization of Differences between Expert and Beginner -, Journal of the Visualization Society of Japan 41, Vol.1, p52-55, 2021(in Japanese)
24. Yun, H., Nakamura, E., Kageyama, Y., Ishizawa, C., Kato, N., Igarashi, K., Suzuki, M.: Basic Study on Development of Action Recognition Technology Based on Human-Skeleton Data, 63rd Japan Joint Automatic Control Conference, 2020(in Japanese)
25. Shimizu, S., Hirayu, H., Asai, H.: Development of Failsafe Camera System for Fastening Work in Manufacturing Scene, The Institute of Electrical Engineers of Japan, Vol.129, No.5, p.916-922, 2009(in Japanese)
26. Kitsukawa, T., Moro, A., Pathak, S., Umeda, K., Harada, Y., Nishikawa, H., Noguchi, M., Hamaya, A.: Progress Estimation of Assembly Work Using Deep Metric Learning, Autumn conference of the Japan Society for Precision Engineering, 2022
27. Murai, K., Imai, T., Arai, K., Kobayashi, T.: Real-time Work Procedure Monitoring System by Deep Learning and Web Camera Using Results of Work Analysis at Manufacturing, Information Processing Society of Japan, CDS, Vol.10, No.3, 1-10, 2020(in Japanese)
28. Yoshii, T., Oshima, H., Shirakawa, Y., Nishimura, K., Kato, T., Maekawa, T. and Namioka, Y.: Human activity recognition in factory based on estimation of product state and human-object interaction area, Proceedings of the 84th National Convention of IPSJ, pp.21-22, 2022(in Japanese)
29. Oshima, H., Shirakawa, Y., Yoshii, T., Kato, T., Maekawa, T. and Namioka, Y.: Automatic Work Progress Recognition based on Product Status Estimation Using Deep Learning, Proceedings of the Manufacturing Systems Division Conference 2022 of JSME, pp.25-30, 2022(in Japanese)
30. Shirakawa, Y., Oshima, H., Yoshii, T., Nishimura, K., Kato, T., Maekawa, T. and Namioka, Y.: Human-object interaction area estimation based on human pose and optical flow in factory operation video, Proceedings of the 84th National Convention of IPSJ, pp.71-72, 2022(in Japanese)
31. Yan, S., Xiong, Y. and Lin, D.: Spatial Temporal Graph Convolutional Networks for Skeleton Based Action Recognition, Proceedings of the Thirty-Second AAAI Conference on Artificial Intelligence, pp.7444-7452(2018).
32. Li, C., Cui, Z., Zheng, W., Xu, C., Ji, R. and Yang, J.: Action-Attending Graphic Neural Network, IEEE Transactions on Image Processing, Vol. 27, pp.3657-3670(2018).
33. Ibsa K, J., Thanh-Dat, T., Jessica L, A., Han-Seok, S. and Khoa, L.: Movement Analysis for Neurological and Musculoskeletal Disorders Using Graph Convolutional Neural Network, MDPI, Future Internet, Vol. 13, No. 8(2021)..