

実倉庫環境における IMU データを用いた作業認識への 自己教師あり学習の適用と有効性評価

渡邊 企章^{1,a)} 加納 一馬¹ ホサイン タヘラ¹ 浦野 健太¹ 米澤 拓郎¹ 河口 信夫^{1,2}

受付日 2025年12月16日, 採録日 2026年7月1日

概要: 産業現場における人の作業内容のデータ化・分析は、業務効率の評価や改善を目的としたデジタルツインの実現において重要な要素である。なかでも、スマートフォンやウェアラブルデバイスに搭載された IMU センサを用いた行動認識技術は、追加設備を必要としない作業認識手法として注目されている。一方で、産業現場の作業は複数の動作が時間的に連続する複合的な行動であるため、これをセンサデータから自動認識するのは容易ではなく、高精度な認識には大規模なラベル付きデータが必要とされている。しかし、実環境におけるアノテーションコストは高く、これが作業認識技術の実用化を妨げる要因となっている。このような課題に対し、本研究ではラベルなしセンサデータから有用な特徴表現を学習可能な自己教師あり学習に着目し、必要なラベル量の削減を目的としてその適用可能性を調査する。具体的には、自己教師あり学習を取り入れた作業認識フレームワークを構成し、その有効性を認識性能、作業特性、および訓練データ量との関係の観点から分析した。物流倉庫における独自の実作業データセットおよび OpenPack データセットを用いた評価実験では、自己教師あり学習により事前学習した CNN と長期時系列モデリングを行う Transformer Encoder を組み合わせた作業認識モデルがベースラインを一貫して上回る認識性能が確認され、作業認識における自己教師あり学習やラベルなしセンサデータ活用の有効性が示された。

キーワード: 行動認識, 自己教師あり学習, Transformer, ウェアラブルセンサ, IMU

Applying Self-supervised Learning to IMU-based Task Recognition in Real-world Warehouse Environments

KISHO WATANABE^{1,a)} KAZUMA KANO¹ TAHERA HOSSAIN¹ KENTA URANO¹ TAKURO YONEZAWA¹
NOBUO KAWAGUCHI^{1,2}

Received: December 16, 2025, Accepted: July 1, 2026

Abstract: In industrial environments, acquiring and analyzing data on human work activities is essential for digital twin applications that aim to evaluate and improve operational efficiency. Activity recognition using inertial sensors embedded in smartphones and wearable devices has therefore attracted attention as a low-cost and easily deployable approach. However, work activities consist of multiple actions occurring sequentially over time, making automatic recognition from sensor data challenging. High recognition accuracy typically requires large amounts of labeled data, but the high annotation cost in real-world environments hinders practical deployment. In this study, we investigate the use of self-supervised learning for task recognition using IMU sensor data. We apply a task recognition framework designed for long-duration work activities and analyze its effectiveness in terms of recognition performance, task characteristics, and training data size. The framework combines self-supervised pretraining for short-term motion representation with a Transformer-based model for capturing long-term temporal context. Experiments on a real-world logistics warehouse dataset and the OpenPack dataset show that the framework achieves improved performance compared to baseline methods under various conditions. The results also demonstrate the effectiveness of leveraging unlabeled sensor data for task recognition and highlight its applicability to real-world industrial environments.

Keywords: human activity recognition, self-supervised learning, Transformer, wearable sensors, IMU

1. はじめに

近年、製造業や物流業では DX 化が進み、人・設備・モノの状態をデータ化して業務効率の分析や改善につなげる「デジタルツイン」に関する取り組みが拡大している [1], [2]. また、全体最適化の実現のために、設備や荷物だけでなく、「人」のデータ活用も注目されている. たとえば、物流倉庫で人の移動軌跡を可視化し、動線の統計分析やシミュレーションに活用する研究が進められている [3], [4], [5]. また、人の作業内容をデータ化し、効率評価や安全管理に役立てる試みも行われている [6]. 特に、スマートフォンやウェアラブルセンサに搭載された IMU を用いた認識手法は追加設備を必要とせず、人の動作を詳細にとらえられるため、行動を自動計測するにあたって有望なアプローチである [7]. ただし、IMU による行動認識は通常、分類モデルを用いて数秒単位のセンサデータから動作を認識する枠組みが中心であり、複数の動作を含む長期的な行動である実作業の認識は容易ではない. そのため、高精度な認識モデルの構築には大規模なラベル付きデータが必要とされているが、現場ごとに認識対象のニーズに合わせたラベル付きデータを大量に用意するのは現実的ではなく、このアノテーションコストが作業認識技術実用化の妨げとなっている. たとえば、産業領域における大規模マルチモーダルデータセットである OpenPack データセット [8] では、1 セッション（20 回の梱包作業・約 30 分）のデータに対し、アノテーション作業だけで約 5 時間を要したと報告されている. また、データセット全体（約 54 時間）の整備については複数名の専門アノテータが約 1 年にわたり作業を行ったと言及されており、ラベル付きデータの大規模収集が容易でないと分かる. このような課題に対し、近年では半自動アノテーション手法 [9] や、自己教師あり学習・ドメイン適応に基づく少量ラベル下での認識手法 [10], [11] が活発に研究されている. 特に自己教師あり学習を用いた手法は、ラベルのないデータから個人差や環境変動に頑健な特徴を獲得し、限られたラベルでも高精度な識別を可能にする点で有望である.

本研究では、産業現場における作業認識の具体例として、物流倉庫における人手作業を対象に扱う. 物流倉庫では、図 1 に示すように入荷から格納に至るまでの工程において、検品、仕分け、搬送といった自動化が難しい作業が人手で行われている. 現状は経験則によって作業シフトやレイアウトが設定されていることが多いが、これらを定量的



図 1 物流倉庫における実作業の例

Fig. 1 Example tasks in a logistics warehouse.

に評価して最適化していくためには、人の作業内容のデータ化が重要な要素となる.

このような実作業における動作の多様性やラベル収集コストといった課題に対し、本研究では自己教師あり学習と時系列モデリングを組み合わせた構成を用いて、センサベースの作業認識における自己教師あり学習を用いたラベルなしセンサデータ活用の有効性を評価する. 評価では、物流倉庫において収集した実データを用いて実環境での性能を検証し、さらに公開データセットを用いて汎用性についても検討する. 本研究の主な貢献は以下のとおりである.

- センサベースの行動認識における自己教師あり学習と時系列モデリングを組み合わせ、短期的な動作特徴と長期的な作業文脈を分離して扱う構成を想定し、産業現場における作業認識への適用を検証した.
- 実際の物流倉庫における実作業データおよび Open-Pack データセットを用いて、自己教師あり学習を活用した作業認識手法の性能を評価した.
- ラベル付きデータ量や作業特性、データセットの違いに着目した分析を行い、自己教師あり学習の効果が条件に応じて変化することを明らかにした.

2. 関連研究

2.1 センサベースの行動認識

スマートフォンに内蔵された加速度センサや角速度センサを用いた行動認識 (Human Activity Recognition, HAR) は、歩行や座位、階段昇降といった日常的な動作の認識に広く用いられてきた [7], [12]. 初期の研究では、平均や分散などの統計量を選定し、SVM や GBDT といった機械学習手法で分類するアプローチが主流であった [12].

近年では、CNN, LSTM, Transformer などの深層学習モデルにより、センサデータから直接特徴量を抽出する end-to-end 型の手法が広く研究されている [13], [14]. これらの手法は日常生活だけでなく、物流や製造業などの産業現場においても、作業記録、業務分析、安全管理などの目的で活用が進んでいる [6]. しかし、産業現場における実作業認識は従来の行動認識が主に対象としてきた「歩行」「座位」といった短時間で完結する動作と異なり、複数の動作の連続で構成されるため、短時間のセンサ系列を入力として行動を認識する枠組みをそのまま適用することは難し

¹ 名古屋大学大学院工学研究科
Graduate School of Engineering, Nagoya University, Nagoya,
Aichi 464-8603, Japan

² 名古屋大学未来社会創造機構
Institutes of Innovation for Future Society, Nagoya University,
Nagoya, Aichi 464-8601, Japan

^{a)} kisho@ucl.nuee.nagoya-u.ac.jp

い。作業の識別には、動作のならびや遷移といった長期的な時系列文脈をとらえる必要があるが、長期依存を扱えるモデルを安定して学習させるには、単一動作の認識以上に多くのラベル付きデータが求められる。しかし、産業現場における作業データのアノテーションは時間的なコストが大きく [8]、十分なデータ量の確保が実運用上の大きな障壁となっている。

2.2 自己教師あり学習（表現学習）の活用

2.2.1 代表的な手法と研究動向

ラベル付きデータの収集コストを低減する手法として、センサベースの行動認識分野では自己教師あり学習（Self-Supervised Learning, SSL）が広く用いられている [15]。自己教師あり学習はラベルなしセンサデータを用いてあらかじめ設定した擬似タスクを解く中で、後の行動認識に有用な特徴表現を獲得する手法である。学習によって得られた特徴表現は、少量のラベル付きデータを用いて軽量の分類ヘッドを学習し、分類などの下流タスクへ転移できる。

自己教師あり学習の代表的な手法として、いくつかの枠組みが提案されている。補助タスクに基づく手法として、Multi-task SSL [16] は、スケーリングやノイズ付加といったデータ拡張が入力信号に適用されたかを識別する複数の補助タスクを同時に学習し、センサデータに共通する有用な特徴表現を獲得する。再構成に基づく手法として、LIMU-BERT [17] は、時系列信号の一部をマスクし、欠損部分を復元するタスクを通じて表現学習を行う。このようなマスク信号復元に基づく学習により、時系列データの周期性や時間的な構造をとらえることが可能となる。一方、対照学習に基づく手法は、センサベースの行動認識分野において特に活発に研究されている。SimCLR [18] や CSSHAR [19] は、コンピュータビジョン分野で提案された SimCLR [20] をセンサデータに応用した手法であり、同一の入力信号に対して異なるデータ拡張を施したペアを正例として特徴を近づけるように学習することで、装着条件や個人差に頑健な動作表現が得られることが報告されている。対照学習の学習戦略は多様であり、負例不要な BYOL [21] や SimSiam [22]、プロトタイプベースの SwAV [23]、正則化ベースの VICReg [24] など、センサデータへの応用を含めて様々な手法が検討されている [15]。

2.2.2 作業認識への適用における課題

産業現場における作業認識では、複数の動作が連続する長期的な文脈をとらえながら、作業員間のばらつきに依存しないロバストな特徴表現を得ることが重要である。しかし、既存の自己教師あり学習手法では、この2つの要件を同時に満たすことは容易ではない。

短期的な動作認識を対象とした既存研究では、自己教師あり学習によって歩行や静止といった短時間の動作ごとに異なる特徴表現を形成できると報告されており、SSL の有

効性が示されている [16], [17]。一方で、これらの手法はいずれも数秒程度の短期ウィンドウを対象として設計されており、ウィンドウを長時間化するだけでは動作遷移や作業手順といった長期的な文脈情報が適切に反映されるとは限らない。たとえば対照学習型手法では局所的な不変性が強調されるため長期構造の抽出が難しく、再構成型手法はセンサのノイズや局所的変動の予測にモデル容量を割くため [15]、長期的な作業プロセスの構造的特徴をとらえにくい。

長期的なセンサパターンを扱う試みとして MoIL [25] が提案されており、反復的な動作パターン（Motif）に着目して中長期の構造を特徴表現として抽出できることが示されている。一方で、同一の作業であっても、扱うアイテムや作業員の違いによって動作の長さや現れ方が大きく変動することが指摘されている。

以上より、SSL は短期動作の認識には有効であるものの、長期的な文脈の理解とユーザロバスト性を同時に満たした特徴表現をラベルなしデータから直接獲得することは容易ではない。そこで本研究では、長期構造とユーザロバスト性を同時に学習させるのではなく、自己教師あり学習によるロバストな基礎動作表現の獲得と、教師あり学習による作業文脈のモデリングを段階的に組み合わせた構成を用いて、作業認識における自己教師あり学習の活用について検討する。

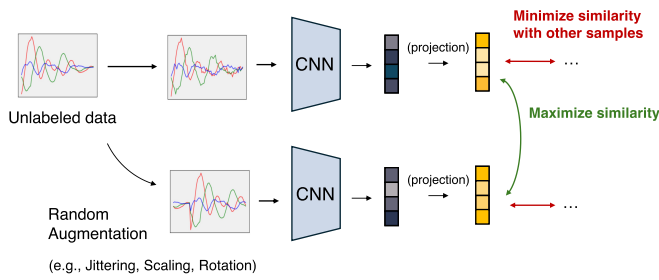
3. 作業認識手法

3.1 学習枠組みの概要

本研究では、長期的な作業認識モデルにおいてラベルなしセンサデータを活用するため、短時間の基本動作に関する特徴抽出と、長期的な時系列文脈のモデリングを組み合わせた構成を用いる。具体的には、作業を短時間に観測される基本動作の系列としてとらえ、その時間的な構造から作業を推定する認識モデルを用いて、短期の特徴抽出器のみを自己教師あり学習により事前学習する。

本研究で用いる作業認識手法の概要を図 2 に示す。第 1 段階では、ラベルなしのセンサデータを用いた自己教師あり学習により、一般化可能な基本動作をとらえる CNN を事前学習する。第 2 段階では、ラベル付きデータを用いて Transformer Encoder, MLP を含む長期時系列分類モデルを学習する。この際、CNN の重みは固定（もしくは fine-tuning）する。この学習により、基本動作の時系列と作業との対応関係を効果的に学習することを目指す。本稿では、このような事前訓練済み CNN と Transformer Encoder からなる作業認識モデルの構成を PreCNN-Transformer (PreCNN-T) と表記する。以上の構成を用いて、産業現場の作業認識におけるラベルなしセンサデータを用いた自己教師あり学習の活用について検討する。

1. 短期特徴抽出器の自己教師あり学習



2. 作業認識モデルの教師あり学習

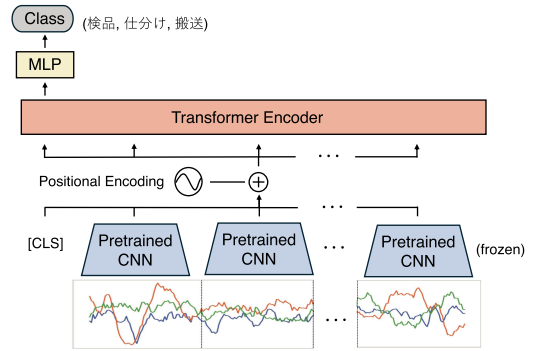


図 2 特徴抽出器の自己教師あり学習を活用した長期作業認識モデル (PreCNN-T)

Fig. 2 Long-term task recognition model using self-supervised learning (PreCNN-T).

表 1 SimCLR による事前学習に用いたデータ拡張の一覧

Table 1 Data augmentations for SimCLR pretraining.

Augmentation (Prob.)	Purpose	Application Unit	Method
ノイズ追加 (1.0)	センサノイズの再現	Channel-wise	平均 0 のガウスノイズを追加 ($\text{std} = 0.05\sigma_s$)
スケーリング (0.5)	動作強度・感度差	Channel-wise	係数を乗じて振幅を変化 ($\sim \mathcal{N}(1.0, 0.7\sigma_s)$)
回転 (0.5)	センサ装着方向のずれ	Synchronized	Acc/Gyro に同一回転行列を適用 (角度 $\pm 30^\circ$)
時間シフト (0.5)	動作開始時刻のずれ	Synchronized	時間軸方向にランダムシフト (最大 2 秒)
時間歪み (0.2)	動作速度のばらつき	Channel-wise	非線形な時間伸縮 (3 次スプライン, 制御点 4)
順序入れ替え (0.2)	局所的順序の変動	Channel-wise	信号の分割と順序入れ替え (分割数 4)

3.2 CNN の自己教師あり学習

3.2.1 SimCLR による自己教師あり学習

本研究では、対照学習に基づく自己教師あり学習の 1 つである SimCLR の枠組みを用いて、ラベルなしのセンサデータから短期的な動作特徴の学習を目指す。

SimCLR に基づく対照学習手法では、ランダムな拡張で生成した 2 つの系列を共通のネットワークで特徴ベクトルに変換し、それらの特徴表現が近くなるように学習を進める。具体的には、ミニバッチ内の各サンプル $\mathbf{x}$ に対して、異なる 2 種類のランダムな変換を適用し、それぞれから得られた 2 つの系列 $\hat{\mathbf{x}}_i, \hat{\mathbf{x}}_j$ を正例ペアとする。この処理によって、バッチサイズが N の場合、 $2N$ 個の変換済みサンプルが得られる。それぞれのサンプルは共通のエンコーダで特徴ベクトル $\mathbf{h}_i, \mathbf{h}_j$ に変換された後、さらに小さな空間へ写像するプロジェクション層を通して、最終的な表現 $\mathbf{z}_i, \mathbf{z}_j$ となる。各正例ペア (i, j) に対して、他の $2N - 2$ 個の埋め込みを負例として扱い、以下の NT-Xent (Normalized Temperature-scaled Cross Entropy) 損失 [20] を最小化し、正例間の距離を縮め、負例とは遠ざけるように学習を行う。

$$\ell(i, j) = -\log \frac{\exp(\text{sim}(\mathbf{z}_i, \mathbf{z}_j)/\tau)}{\sum_{k=1}^{2N} I_{[k \neq i]} \exp(\text{sim}(\mathbf{z}_i, \mathbf{z}_k)/\tau)} \quad (1)$$

ここで、 $\text{sim}(\mathbf{z}_i, \mathbf{z}_j)$ は通常、コサイン類似度によって定義される。 τ はスケーリング係数として働き、値が小さいほど類似度の違いを強調する。また、 $I_{[k \neq i]}$ はインデ

クス $k \neq i$ のときに 1, $k = i$ のときに 0 をとる指示関数 (Indicator Function) である。この損失により、同一入力に由来する特徴表現は近づき、異なる入力に由来する表現は遠ざかるように学習が進む。

3.2.2 データ拡張設定

本研究における事前学習の目的は、短時間のセンサ系列から、同一の基本動作に由来する信号が特徴空間上で近接するようなロバストな動作特徴表現を獲得することである。このため、センサに適用するデータ拡張は、実環境で生じうる変動を反映していることが望ましいが、どの拡張が有効かはデータセットに依存することが知られている [19]。そこで本研究では、SimCLR HAR [18] や CSSHAR [19] などのセンサ領域の SimCLR ベース自己教師あり学習や、時系列データ拡張に関する研究 [26] で用いられている拡張の中から、実環境で想定されるセンサ変動を反映する拡張を選定した。適用確率は、対照学習では正例ペアが同一の基本動作を表し続ける必要があることをふまえ、動作の意味を保ちやすい変換ほど高く設定した。本研究で用いた拡張の一覧と適用確率を表 1 に示す。

3.3 作業認識モデルのアーキテクチャ設計

本節では、自己教師あり学習を活用した作業認識モデル (PreCNN-T) の構成について、図 3 を用いて、各構成要素とその処理の流れを説明する。本モデルは、数秒レベルの短時間入力から基本動作表現を抽出する事前学習済み

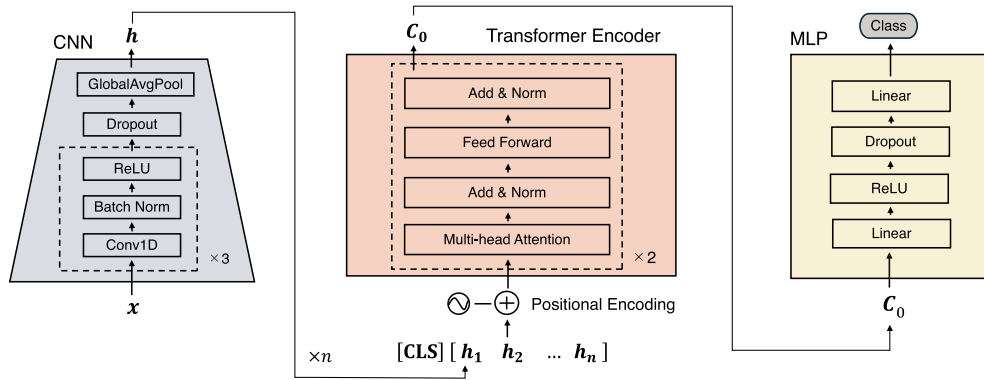


図 3 作業認識モデルのアーキテクチャ

Fig. 3 Architecture of the task recognition model.

CNN と、その動作特徴系列から数十秒スケールの文脈をモデル化する Transformer Encoder、および分類用 MLP から構成される。なお、入力信号のサンプリング条件やウィンドウ長、ならびに CNN の一部のパラメータは、使用するデータセットの特性に応じて設定されるため、以下では、モデル構造の概要と各構成要素の役割について述べる。

入力として、長期センサ時系列 $\mathbf{X} \in \mathbb{R}^{C \times L_{\text{long}}}$ が与えられる。ここで、 C はセンサチャネル数、 L_{long} は系列全体の長さを表す。本研究では、3 軸加速度（重力加速度を含む）および 3 軸角速度を用いるため、 $C = 6$ とする。

入力系列 $\mathbf{X}$ は、一定長の短時間ウィンドウに分割され、各ウィンドウは $\mathbf{x}_t \in \mathbb{R}^{C \times L_{\text{short}}}$ として表現される。各ウィンドウに対して、共通の事前訓練済み CNN を適用することで、固定次元の動作特徴ベクトル $\mathbf{h}_t \in \mathbb{R}^d$ を抽出する。この処理により、長期間のセンサ系列は、 $[\mathbf{h}_1, \mathbf{h}_2, \dots, \mathbf{h}_T]$ からなる動作特徴系列へと変換される。ここで、 T はウィンドウ分割後の系列長を表す。

得られた動作特徴系列の先頭には分類用の [CLS] トークンを付加し、Positional Encoding を加えた後、Transformer Encoder に入力する。Transformer Encoder では自己注意機構を用いて動作間の時間的関係を考慮し、作業全体を表す文脈表現を学習する。

Transformer Encoder から得られる [CLS] トークンに対応する出力 $\mathbf{c}_0 \in \mathbb{R}^d$ は、作業全体を要約した文脈ベクトルとして解釈される。この文脈ベクトルを入力として MLP による分類を行い、作業クラスを推定する。

4. 評価

本研究では、スマートフォンの IMU データを用いた産業現場作業認識に対する自己教師あり学習の有効性を検証するため、実環境で収集したデータを用いた評価実験を行う。まず、物流倉庫における実作業データを対象として、短期および長期の作業認識モデルを構築し、自己教師あり学習の有無を含めて、(1) 被験者単位交差検証による認識性能、(2) 作業特性と認識精度の関係、(3) 訓練データ量と

認識精度の観点から評価する。さらに、作業構造や計測条件が異なる他の公開データセットを用いた検証も行う。

4.1 データセット

4.1.1 物流倉庫実作業データセット

本研究では、物流倉庫入荷場における実作業に関するデータを収集し、限られたラベル量での作業認識について検証する。データ収集の際は、物流倉庫の実際の作業員に、腰部にスマートフォンを装着した状態で通常業務を行ってもらい、その間の 3 軸加速度および角速度データを計測し、100 [Hz] にリサンプリングしてデータセットを作成した。1 日の計測において 29 人の作業員からデータを収集し、合計 196 時間分のセンサデータを取得した。このうち、一部の被験者については作業中の様子をカメラ映像を参照しながら各時刻に対応する実施作業を人手でラベル付けして、正解データを作成した。ラベル付きデータは 10 人分、合計 6 時間である。

対象とする作業ラベルは、商品が入荷されてから倉庫に格納されるまでの以下の人手作業である。

- 検品 (Inspect)：商品と注文書を照合する作業
- 仕分け (Sort)：商品を適切な場所に振り分ける作業
- 搬送 (Transport)：荷物を載せた台車を運搬する作業

これらの作業は一定の役割分担が存在するものの、状況に応じて柔軟に交替され、作業が数分で切り替わる場合もあれば、数十分継続する場合もある。ラベル付きデータにおける各作業の出現割合は、図 5 に示すとおりである。

4.1.2 OpenPack データセット

多種多様な作業認識における自己教師あり学習の有効性や特性をより詳細に確認するため、公開データセットである OpenPack データセット [8] を用いた評価も行う。OpenPack データセットは、物流倉庫における梱包作業を対象として収集されたマルチモーダル行動認識データセットであり、Picking や Assemble Box などの 10 種類の作業とそれらを構成するより詳細な 32 種類の動作に関するラベル付きデータが 16 人の被験者から合計で約 54 時間分提

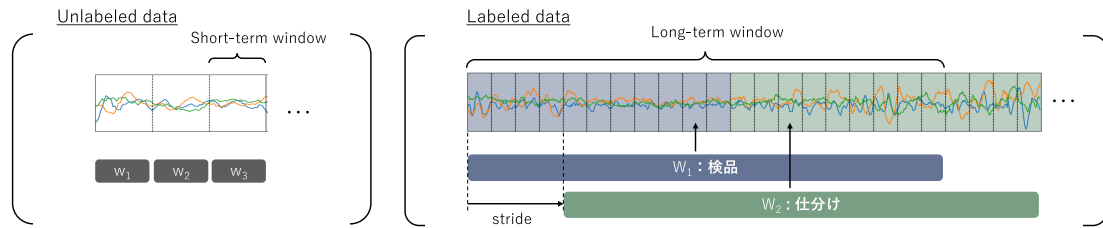


図 4 センサデータのウィンドウ処理

Fig. 4 Windowing process for sensor data.

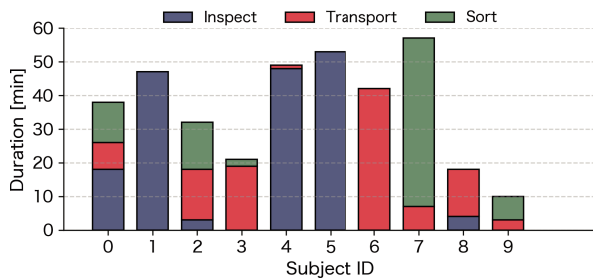


図 5 正解データにおけるラベル内訳

Fig. 5 Ground-truth label distribution.

供されている。これらの作業は、数秒から数十秒程度の時間スケールで構成されており、本研究で対象とする作業認識問題と同様に、複数の動作が連続して構成される時系列構造を有する。評価実験では、OpenPack データセットに含まれる左手首の IMU データを使用し、10 種類の作業を認識するモデルを構築する。

4.2 データのウィンドウ処理

収集したセンサデータに施すウィンドウ処理の流れを、図 4 に示す。まず、CNN の事前学習に用いるラベルなしセンサデータに対しては、長時間の時系列データを、CNN の入力長に対応する短時間ウィンドウ単位に分割し、これらを入力として SimCLR による自己教師あり学習を行う。

一方、作業認識モデルの訓練に用いるラベル付きセンサデータに対しては、長時間の時系列データを、作業認識モデルへの入力長に対応するウィンドウ単位に分割し、これを一定のストライドで時系列方向にスライドさせながら切り出す。各ウィンドウには、その中央時刻に対応する作業ラベルを付与し、作業認識のための教師データを構築する。

本実験では、短時間ウィンドウのサイズを 4 [s]、長時間ウィンドウのサイズを 40 [s] とした。また、推論時・学習時のストライドはどちらも 4 [s] に設定した。

4.3 比較手法

自己教師あり学習やモデル構成の違いが認識性能に与える影響を評価するため、入力長、事前学習の有無、および時系列モデリング手法の違いに基づき、複数の比較手法を設定する。表 2 に本実験で用いる認識モデルの一覧を示す。

CNN および PreCNN は、短時間のセンサ系列を入力と

表 2 比較対象の認識モデル

Table 2 Recognition models for comparison.

Model	Pretraining	Input	Temporal model
CNN	–	Short (4 s)	–
PreCNN	✓	Short (4 s)	–
DCL	–	Long (40 s)	LSTM
CNN-T	–	Long (40 s)	Transformer
PreCNN-T	✓	Long (40 s)	Transformer

し、CNN によって動作特徴を抽出した後、MLP により分類を行うベースラインとして用いる。PreCNN では、CNN に対して SimCLR による事前学習を適用する。

DeepConvLSTM (DCL) [13] は、CNN による特徴抽出と LSTM による時系列モデリングを組み合わせた代表的な行動認識モデルであり、長期依存関係を考慮する深層学習ベースラインとして採用する。

CNN-T および PreCNN-T は、長時間のセンサ系列を入力とし、CNN による特徴抽出と Transformer Encoder による長的文脈モデリングを組み合わせた作業認識モデルである。PreCNN-T は CNN 部分を SimCLR で事前学習した認識モデルである。また、PreCNN、PreCNN-T については、SimCLR におけるデータ拡張設定として 2 つの設定を比較する。1 つは、CSSHAR [19] が UCI-HAR データセットにおいて最良と報告した 3 種の拡張（ノイズ追加、スケーリング、順序入れ替え）を用いる設定であり、もう 1 つは、6 種の拡張（ノイズ追加、スケーリング、回転、時間シフト、時間歪み、順序入れ替え）を用いる設定である。

4.4 交差検証による評価

モデルの認識精度の評価は、被験者単位での 1 人抜き交差検証 (Leave-One-Subject-Out Cross Validation, LOSO-CV) により行う。具体的には、各被験者を 1 名ずつテストデータとし、残りの被験者を訓練データとしてモデルを学習・評価する。この手順を全被験者について繰り返し、モデルの認識精度や汎化性能を評価する。

物流倉庫実作業データセットでは、190 時間分のラベルなしデータを用いて SimCLR による CNN の事前訓練を行う。その後、10 人分・計 6 時間のラベル付きデータを用いて、被験者間交差検証による作業認識性能を評価する。

OpenPack データセットでは、約 40 時間分のラベルなし

表 3 CNN のアーキテクチャ
Table 3 CNN architecture.

Component	Configuration
Input	IMU ($C \times L_{\text{short}}$)
Conv1	$C \rightarrow 64$, kernel = 10, stride = 4
Conv2	$64 \rightarrow 64$, kernel = 6, stride = 2
Conv3	$64 \rightarrow 128$, kernel = 4, stride = 2
Normalization	Batch Normalization
Activation function	ReLU
Pooling	Global Average Pooling
Dropout rate	0.3
Output	Feature vector (128-dim)

表 4 Transformer encoder のアーキテクチャ
Table 4 Transformer Encoder architecture.

Component	Configuration
Input	Feature sequence ($W_{\text{num}} \times 128$)
Embedding dimension	128
Number of layers	3
Attention heads	4
Feed-forward dimension	256
Positional encoding	Sinusoidal
Dropout rate	0.1
Output	context vector (128-dim)

センサデータを用いて CNN の事前訓練を行い、16 人分・計 10 時間のラベル付きデータを用いて被験者間交差検証による評価を行う。

さらに、訓練データ量と認識精度の関係を評価するため、交差検証における訓練データの各被験者・各クラスに対して、1%, 10%, 25%, 50%, 75%, 100% の割合でサンプリングを行い、クラス比率を維持した訓練データを構成し、それぞれの条件での認識精度を評価した。なお、サンプリング時は各クラスで最低 1 サンプルを確保している。

評価指標には F1 スコアを用いる。特に物流倉庫実作業データセットはクラス間の偏りが大きく、被験者ごとに一部クラスのサンプル数が極端に少ない場合があるため、各クラスを等しく扱う Macro-F1 スコアに加えて、各クラスのサンプル数に基づく重み付き平均である Weighted-F1 スコアを採用し、全体の認識性能の指標としている。

4.5 実験の詳細

4.5.1 モデル設計

自己教師あり学習を活用した作業認識モデル PreCNN-T は、短時間ウィンドウごとに適用される CNN、それらの系列から長期的な作業文脈をとらえる Transformer Encoder、および最終的な作業分類を行う MLP から構成される。PreCNN-T のアーキテクチャ設計を表 3、表 4、表 5 に示す。比較手法は、これらの構成要素の一部を用いており、ハイパーパラメータを共有している。

表 5 MLP classifier のアーキテクチャ
Table 5 MLP classifier architecture.

Component	Configuration
Input	context vector (128-dim)
Linear 1	$128 \rightarrow 128$
Linear 2	$128 \rightarrow K$
Activation function	ReLU
Dropout rate	0.3
Output	K classes

4.5.2 学習設定

本節では、各モデルに共通する学習設定および最適化条件について述べる。評価は被験者単位の交差検証に基づいて行い、各分割において、1 人分を評価用、残りを訓練用とし、さらに訓練データの 10% を検証用データとしてハイパーパラメータ調整および early stopping に利用した。

CNN の事前学習には SimCLR を用い、ラベルなしセンサデータを用いて自己教師あり学習を行った。拡張の種類と適用確率として、3 種類の拡張を用いる設定では、ノイズ追加 (1.0)、スケーリング、順序入れ替え (各 0.5) を適用し、6 種類の拡張を用いる設定では、ノイズ追加 (1.0)、スケーリング、回転、時間シフト (各 0.5)、順序入れ替え、時間歪み (各 0.2) を加えた拡張を適用した。損失関数には NT-Xent 損失を用い、バッチサイズ 256、学習率 1.0×10^{-3} の設定のもと、Adam [27] 最適化により最大 100 エポック訓練した。事前学習によって得られた CNN は、作業認識モデルにおいて特徴抽出器として利用する。

作業認識モデルの学習では、CNN を事前学習しない場合は、CNN をランダム初期化したうえでネットワーク全体をエンドツーエンドで学習させた。一方、事前学習済み CNN を用いる場合は、CNN の重みを固定し、後続層のみを学習させた。分類損失としてラベル不均衡を考慮した重み付き交差エントロピー損失を用いた。すべての認識モデルにおいて、バッチサイズ 16、学習率 1.0×10^{-3} 、最大エポック数 30 の設定で Adam により最適化を行った。検証損失が連続して改善しない場合には early stopping を適用した。

4.6 評価結果

4.6.1 物流倉庫実作業データセット

物流倉庫実作業データセットに対して、被験者間交差検証を行い、各モデルの認識性能を評価した。表 6 に各モデルについて全体の認識性能を示す。PreCNN および PreCNN-T は、事前学習をしない場合と比べて Macro-F1、Weighted-F1 の両指標が向上しており、特に拡張 6 種の PreCNN-T は両指標において最もスコアが高く、Macro-F1=0.502、Weighted-F1=0.852 で 3 つの作業を分類した。また、拡張 6 種の設定は 3 種の場合と比較して一貫して高い性能を示した。以降では、PreCNN、PreCNN-T は拡張

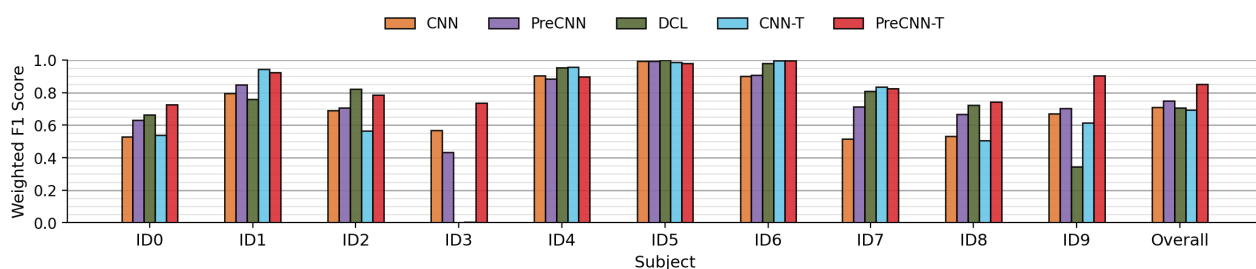


図 6 被験者ごと F1 スコアの分布 (物流倉庫実作業データセット)

Fig. 6 Subject-wise F1 scores on the warehouse dataset.

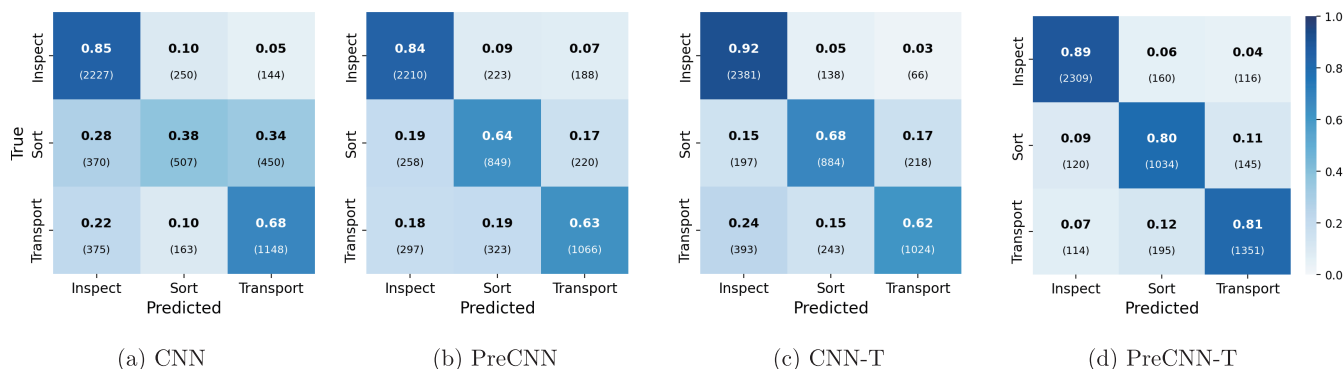


図 7 混同行列の比較 (物流倉庫実作業データセット)

Fig. 7 Confusion matrices on the warehouse dataset.

表 6 全体の認識スコア (物流倉庫実作業データセット)

Table 6 Overall recognition scores on the warehouse dataset.

Model	Macro-F1	Weighted-F1
CNN	0.375	0.710
PreCNN (拡張 3 種)	0.380	0.727
PreCNN (拡張 6 種)	0.405	0.749
DCL	0.420	0.705
CNN-T	0.363	0.694
PreCNN-T (拡張 3 種)	0.454	0.793
PreCNN-T (拡張 6 種)	0.502	0.852

6 種で訓練されたモデルを指す。なお、本データセットでは作業クラス間の出現頻度の偏りが大きく、少数クラスの誤認識が Macro-F1 の全体指標を下げ、両指標の値に差が生じている。

図 6 は、被験者ごとの Weighted-F1 スコアの分布を示したものである。長期時系列モデルを用いた手法では高い認識性能が得られる場合もあるが、被験者ごとの性能変動が大きい傾向が確認できる。これに対し、事前学習を導入した PreCNN-T では、被験者間のばらつきが相対的に小さく、安定した作業認識が実現されている。

図 7 に各モデルの混同行列を示す。CNN 単体では、「仕分け (Sort)」と「搬送 (Transport)」の間で顕著な誤分類が見られ、短時間の動作特徴のみでは両作業を十分に区別できていないことが分かる。これに対し、PreCNN (図 7(b)) では、仕分けと搬送作業の混同が低減した。また、Transformer による時系列モデリングを導入した CNN-T

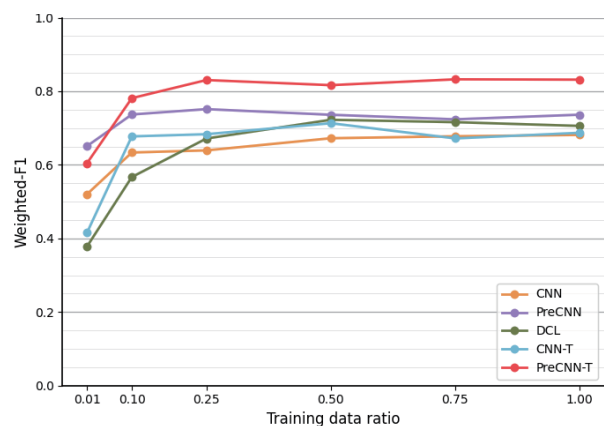


図 8 訓練データ量と認識精度の関係 (物流倉庫実作業データセット)

Fig. 8 F1 scores at different training data ratios on the warehouse dataset.

(図 7(c)) においても、作業文脈の考慮により誤分類が減少している。さらに、事前学習と時系列モデリングを組み合わせた PreCNN-T (図 7(d)) では、仕分けおよび搬送の混同が大きく改善し、双方において最も高い識別精度が得られている。

また、図 8 には訓練データ量と認識精度の関係を示す。自己教師あり学習を導入したモデルは、少量の訓練データにおいても比較的高い性能を示しており、特に PreCNN-T は、PreCNN と比較して訓練データ量の増加にともなう性能改善の度合いが大きく、より高い性能に到達する傾向が見られた。

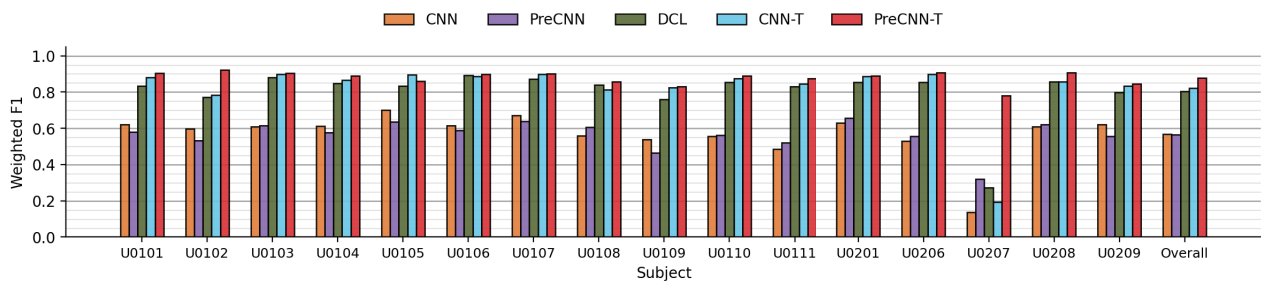


図 9 被験者ごと F1 スコアの分布 (OpenPack データセット)

Fig. 9 Subject-wise F1 scores on OpenPack.

表 7 全体の認識スコア (OpenPack データセット)

Table 7 Overall recognition scores on OpenPack.

Model	Macro-F1	Weighted-F1
CNN	0.568	0.569
PreCNN (拡張 3 種)	0.505	0.506
PreCNN (拡張 6 種)	0.562	0.564
DCL	0.780	0.803
CNN-T	0.801	0.821
PreCNN-T (拡張 3 種)	0.805	0.823
PreCNN-T (拡張 6 種)	0.864	0.879

表 8 作業の持続時間と F1 スコアの比較 (OpenPack データセット)

Table 8 Task durations and F1 scores on OpenPack.

Label	Duration [s]	F1-score	
		CNN-T	PreCNN-T
Assemble Box	20.3 ± 7.0	0.87	0.91
Relocate Item Label	18.2 ± 12.8	0.82	0.90
Scan Label	17.0 ± 9.2	0.82	0.87
Close Box	13.2 ± 4.0	0.83	0.89
Picking	12.3 ± 6.0	0.80	0.87
Insert Items	11.7 ± 6.1	0.82	0.85
Fill out Order	10.1 ± 4.6	0.81	0.86
Attach Shipping Label	9.1 ± 3.0	0.84	0.88
Put on Back Table	6.0 ± 2.1	0.77	0.83
Attach Box Label	5.8 ± 2.3	0.64	0.76

4.6.2 OpenPack データセット

OpenPack データセットを用いて被験者間交差検証による作業認識性能評価を行った。表 7 に各モデルの全体性能を示す。本データセットにおいても、PreCNN-T が最も高い認識性能を示し、Macro-F1=0.864, Weighted-F1=0.879 で 10 個の作業を分類した。また、被験者ごとの Weighted-F1 スコアの分布を示す図 9 から、PreCNN-T は多くの被験者に対して安定して高い F1 スコアを示していることが分かる。

また、本実験で用いたデータにおける作業ごとの持続時間に関する統計情報と F1 スコアの関係をまとめたものを表 8 に示す。まず、CNN-T では、Assemble Box (20.3 ± 7.0 [s]) や Relocate Item Label (18.2 ± 12.8 [s]) といった比較的長時間かつ持続時間のばらつきが大きい作業においても、F1 スコア 0.8 以上で分類できている。一方で、

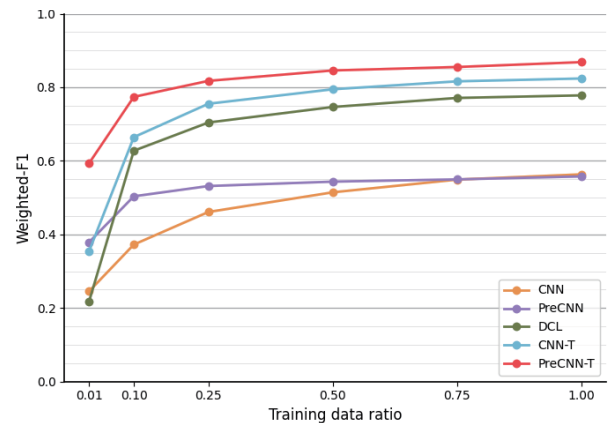


図 10 訓練データ量と認識精度の関係 (OpenPack データセット)

Fig. 10 F1 scores at different training data ratios on OpenPack.

Attach Box Label (5.8 ± 2.3 [s]) のような短時間かつ局所的な動作から構成される作業では認識性能が相対的に低くなる傾向が見られた。また、自己教師あり学習を導入した PreCNN-T では同様の傾向が見られつつも、すべての作業において F1 スコアが向上しており、全体的な性能改善が確認された。

続いて、図 10 に訓練データ量と認識精度の関係を示す。本データセットにおいても、自己教師あり学習を導入したモデルが訓練データが少ない段階で比較的高い認識性能を示す傾向が確認された。また、モデルのアーキテクチャが同一の場合、訓練データ量が少ない条件ではモデル間の性能差が大きく、データ量の増加にともないその差は縮小する様子が見られた。

さらに、PreCNN-T の実際の作業シーケンスに対する時系列予測の例として、ある被験者 (U0207) における約 5 分間の認識結果を図 11 に示す。認識結果と正解ラベルを比べると、多少のラベル変動が確認されるものの、作業全体の推移という観点では、実際の作業シーケンスと整合した予測結果が得られていた。

4.7 考察

評価結果より、本研究で用いた PreCNN-T は、両データセットにおいて全体性能および被験者ごとの安定性の両面で最も高い性能を示した。また、混同行列の改善や被験者

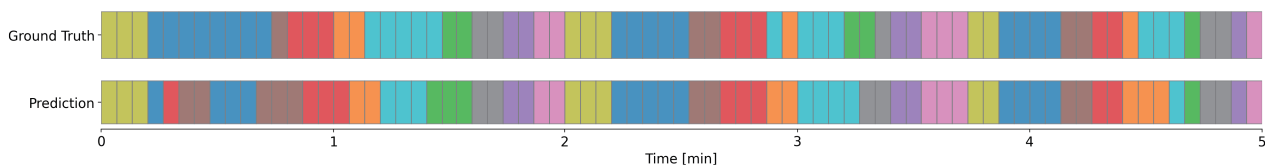


図 11 作業シーケンスに対する認識結果の例 (OpenPack データセット)

Fig. 11 Example recognition results for an OpenPack task sequence.

間ばらつきの低減, さらに訓練データ量が少ない条件における性能の向上など, 複数の観点から性能改善が確認された. 以下では, これらの結果をふまえ, 性能向上の要因について考察する.

まず, 物流倉庫実作業データセットにおける混同行列の分析からは, CNN 単体で顕著であった「仕分け (Sort)」と「搬送 (Transport)」の混同が, 事前学習または時系列モデリングの導入によってそれぞれ緩和され, 両者を組み合わせた PreCNN-T では最も安定した識別が達成されていることが確認できる. 特に, 仕分けや搬送は体全体を用いる作業であり, 作業や状況による動作のばらつきが大きいという, 歩行などの共通動作も含まれるため, 短時間の入力のみでは区別が難しいと考えられる. これに対し, PreCNN-T では, 自己教師あり学習により短期動作に関する頑健な特徴表現を獲得し, さらに時系列モデリングにより長期的な文脈情報を考慮することで, これらの課題がそれぞれ相補的に緩和されたと考察される.

OpenPack データセットにおける作業の持続時間と F1 スコアの関係では, CNN-T は長時間作業では高い性能を示す一方, 短時間作業では性能が低下する傾向が見られた. 一方で, PreCNN-T では特に短時間作業で改善が大きかったが, 平均継続時間が数十秒続くような作業においても一貫した性能向上が見られた. このことから, 自己教師あり学習によって得られる一貫した特徴表現は, 短時間作業の認識だけでなく, 長期時系列モデルの学習の安定化にも寄与していると考えられる. さらに, こうした効果は被験者間にも及んでいる. 一部の被験者で見られた認識精度の大きな低下は, センサ装着状態や個人差による動作のばらつきに起因すると考えられるが, PreCNN-T ではこれが緩和されている. 前述の頑健な動作特徴表現と長期的な文脈のモデリングにより, 個人差の大きい被験者においても作業分類に有効な特徴が得られたと考察される.

次に, 訓練データ量と認識精度の関係では, 訓練データが少ない条件において自己教師あり学習を導入したモデルの性能差が顕著であり, データ量の増加にともないその差は縮小する傾向が確認された. 一方で, 長期時系列モデルでは, データ量が増加した場合でも PreCNN-T が相対的に高い性能を維持していた. この結果は, 特に長期時系列モデルにおいては, 事前学習による個人差に頑健な特徴表現の獲得が性能向上に大きく寄与する可能性を示唆している.

また, SimCLR による自己教師あり学習の際のデータ拡張の設定に関して, 本実験では拡張の種類が多い設定で性能が向上する傾向が確認された. これは, 多様な変換によりセンサ信号のばらつきに対して頑健な特徴表現が学習されたためと考えられる. 一方で, 拡張の種類が少ない場合でも, 事前学習後に作業認識モデルを訓練する際 fine-tuning を行うことで性能が改善される余地がある. これらの関係については, 拡張手法と学習戦略の相互作用の観点から, 今後詳細な検討が必要である.

5. まとめ

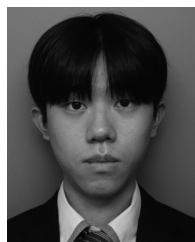
本研究では, 産業現場における人手作業のデータ化という課題に対し, IMU センサを用いた作業認識手法を検討した. 特に, 訓練データ作成時のアノテーションコストを削減するために自己教師あり学習に着目し, その有効性を検証した. 実験では, 短期的な動作特徴の抽出と長期的な作業文脈のモデリングを分離して扱う構成 (PreCNN-T) を用いて, 物流倉庫実作業データセットおよび OpenPack データセットにおける作業認識性能の評価を行った. 評価の結果, PreCNN-T は単純な教師あり学習で訓練される手法を上回る認識性能と安定性を示した. 特に, 特徴抽出器の自己教師あり学習は長期的な時系列モデリングとの組み合わせにより, その効果が顕在化し, 平均持続時間や動作のばらつきが異なる様々な作業に対しても認識精度を向上させることが確認された.

今後は, データ拡張の設計や適用確率が認識性能に与える影響について, より体系的な分析を行うとともに, 拡張手法と fine-tuning を含む学習戦略の最適な組み合わせについて検討する必要がある. また, 本研究では実環境データおよび公開データセットの双方で評価を行ったが, データ特性の違いにより自己教師あり学習による性能改善の現れ方が異なることが確認された. 産業現場における多様な作業や条件においても, 長期時系列モデリングにおける自己教師あり学習の活用が一般化性能の向上に寄与するかについては, より多様なデータセットや異なる作業環境, センサ配置の違いを含めたさらなる検証が必要である.

謝辞 本研究は国立研究開発法人新エネルギー・産業技術総合開発機構 (NEDO) の委託業務 (JPNP23003), 科研費挑戦的研究 (開拓) JP22K18422, 内閣府 SIP3 (JPJ012495), トラスコ中山株式会社に支援いただいている.

参考文献

- [1] Kaiblinger, A. and Woschank, M.: State of the art and future directions of digital twins for production logistics: A systematic literature review, *Applied Sciences*, Vol.12, No.2, p.669 (2022).
- [2] Leng, J., Zhang, H., Yan, D., Liu, Q., Chen, X. and Zhang, D.: Digital twin-driven manufacturing cyber-physical system for parallel controlling of smart workshop, *Journal of Ambient Intelligence and Humanized Computing*, Vol.10, pp.1155–1166 (2019).
- [3] 森 裕輝, 加納一馬, 浅井悠佑, 片山 晋, 浦野健太, 米澤拓郎, 河口信夫: 広角カメラの画像歪みを考慮した物流倉庫内マルチカメラトラッキング手法, 研究報告モバイルコンピューティングと新社会システム (MBL), Vol.2025-MBL-115, No.58, pp.1–8 (2025).
- [4] Kano, K., Mori, Y., Katayama, S., Urano, K., Yonezawa, T. and Kawaguchi, N.: CorVS: Person Identification via Video Trajectory–Sensor Correspondence in a Real-World Warehouse, *2025 International Conference on Indoor Positioning and Indoor Navigation (IPIN)*, pp.1–7 (online), DOI: 10.1109/IPIN66788.2025.11213222 (2025).
- [5] Kawaguchi, N., Asai, Y., Kano, K., Takaki, K., Mori, Y., Suzuki, Y., Watanabe, K., Gushi, Y., Katayama, S., Urano, K., et al.: Digitization Methods for a Logistics Warehouse Towards Digital Twin-Driven Optimization, *2025 IEEE International Conference on Smart Computing (SMARTCOMP)*, pp.186–193, IEEE (2025).
- [6] Niemann, F., Lüdtke, S., Bartelt, C. and Ten Hompel, M.: Context-aware human activity recognition in industrial processes, *Sensors*, Vol.22, No.1, p.134 (2021).
- [7] Straczekiewicz, M., James, P. and Onnela, J.-P.: A systematic review of smartphone-based human activity recognition methods for health research, *NPJ Digital Medicine*, Vol.4, No.1, p.148 (2021).
- [8] Yoshimura, N., Morales, J., Maekawa, T. and Hara, T.: Openpack: A large-scale dataset for recognizing packaging works in iot-enabled logistic environments, *2024 IEEE International Conference on Pervasive Computing and Communications (PerCom)*, pp.90–97, IEEE (2024).
- [9] Higashiura, K., Yokoyama, K., Asai, Y., Shimosato, H., Kano, K., Katayama, S., Urano, K., Yonezawa, T. and Kawaguchi, N.: Semi-automated framework for digitalizing multi-product warehouses with large scale camera arrays, *2024 IEEE International Conference on Pervasive Computing and Communications (PerCom)*, pp.98–105, IEEE (2024).
- [10] Chakma, A., Faridee, A.Z.M., Ghosh, I. and Roy, N.: Domain adaptation for inertial measurement unit-based human activity recognition: A survey, arXiv preprint arXiv:2304.06489 (2023).
- [11] Yuan, H., Chan, S., Creagh, A.P., Tong, C., Acquah, A., Clifton, D.A. and Doherty, A.: Self-supervised learning for human activity recognition using 700,000 person-days of wearable data, *NPJ Digital Medicine*, Vol.7, No.1, p.91 (2024).
- [12] Lara, O.D. and Labrador, M.A.: A survey on human activity recognition using wearable sensors, *IEEE Communications Surveys & Tutorials*, Vol.15, No.3, pp.1192–1209 (2012).
- [13] Ordóñez, F.J. and Roggen, D.: Deep convolutional and lstm recurrent neural networks for multimodal wearable activity recognition, *Sensors*, Vol.16, No.1, p.115 (2016).
- [14] Dirgová Luptáková, I., Kubovčík, M. and Pospíchal, J.: Wearable sensor-based human activity recognition with transformer model, *Sensors*, Vol.22, No.5, p.1911 (2022).
- [15] Logacjov, A.: Self-supervised learning for accelerometer-based human activity recognition: A survey, *Proc. ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, Vol.8, No.4, pp.1–42 (2024).
- [16] Saeed, A., Ozcelebi, T. and Lukkien, J.: Multi-task self-supervised learning for human activity detection, *Proc. ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, Vol.3, No.2, pp.1–30 (2019).
- [17] Xu, H., Zhou, P., Tan, R., Li, M. and Shen, G.: Limubert: Unleashing the potential of unlabeled data for imu sensing applications, *Proc. 19th ACM Conference on Embedded Networked Sensor Systems*, pp.220–233 (2021).
- [18] Tang, C.I., Perez-Pozuelo, I., Spathis, D. and Mascolo, C.: Exploring contrastive learning in human activity recognition for healthcare, arXiv preprint arXiv:2011.11542 (2020).
- [19] Khaertdinov, B., Ghaleb, E. and Asteriadis, S.: Contrastive self-supervised learning for sensor-based human activity recognition, *2021 IEEE International Joint Conference on Biometrics (IJCB)*, pp.1–8, IEEE (2021).
- [20] Chen, T., Kornblith, S., Norouzi, M. and Hinton, G.: A simple framework for contrastive learning of visual representations, *International Conference on Machine Learning*, pp.1597–1607, PMLR (2020).
- [21] Grill, J.-B., Strub, F., Althé, F., Tallec, C., Richemond, P., Buchatskaya, E., Doersch, C., Avila Pires, B., Guo, Z., Gheshlaghi Azar, M., et al.: Bootstrap your own latent-a new approach to self-supervised learning, *Advances in Neural Information Processing Systems*, Vol.33, pp.21271–21284 (2020).
- [22] Chen, X. and He, K.: Exploring simple siamese representation learning, *Proc. IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp.15750–15758 (2021).
- [23] Caron, M., Misra, I., Mairal, J., Goyal, P., Bojanowski, P. and Joulin, A.: Unsupervised learning of visual features by contrasting cluster assignments, *Advances in Neural Information Processing Systems*, Vol.33, pp.9912–9924 (2020).
- [24] Bardes, A., Ponce, J. and LeCun, Y.: Vicreg: Variance-invariance-covariance regularization for self-supervised learning, arXiv preprint arXiv:2105.04906 (2021).
- [25] Xia, Q., Maekawa, T., Morales, J., Hara, T., Oshima, H., Fukuda, M. and Namioka, Y.: Preliminary investigation of ssl for complex work activity recognition in industrial domain via mail, *2024 IEEE International Conference on Pervasive Computing and Communications Workshops and other Affiliated Events (PerCom Workshops)*, pp.465–468, IEEE (2024).
- [26] Iwana, B.K. and Uchida, S.: An empirical survey of data augmentation for time series classification with neural networks, *Plos One*, Vol.16, No.7, p.e0254841 (2021).
- [27] Kingma, D.P. and Ba, J.: Adam: A method for stochastic optimization, arXiv preprint arXiv:1412.6980 (2014).



渡邊 企章 (学生会員)

2025 年名古屋大学工学部電気電子情報工学科卒業。同年より同大学院博士前期課程。産業現場における行動認識、作業データ活用に関する研究に従事。



加納 一馬 (学生会員)

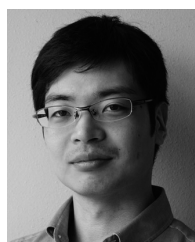
2022 年名古屋大学工学部電気電子情報工学科卒業。2024 年同大学大学院工学研究科情報・通信工学専攻修士課程修了。同年 4 月より同専攻博士課程学生。深層学習を用いた屋内位置推定や画像認識におけるアノテーション効率化に関する研究に従事。



ホサイン タヘラ

2021 年九州工業大学大学院生命体工学研究科人間知能システム工学専攻博士課程修了。ミシガン大学客員研究員、青山学院大学博士研究員、同大学特任助教を経て、2024 年より名古屋大学特任助教。ヘルスケア分野における

ヒューマンコンピュータインタラクション、応用機械学習、マルチモーダルデータ解析に関する研究に従事。ACM, ACM SIGCHI 各会員。



浦野 健太 (正会員)

2016 年名古屋大学工学部電気電子・情報工学科卒業。2018 年同大学大学院修士課程、2021 年博士課程を修了。同年より同大学院助教。博士(工学)。屋内位置推定、実世界データモデリング、生体信号のエンターテインメント

利用に関する研究に従事。



米澤 拓郎 (正会員)

2010 年慶應義塾大学大学院政策・メディア研究科後期課程博士号取得後、同大学院特任助教、特任講師、特任准教授を経て、2019 年より名古屋大学大学院工学研究科准教授。主に、ユビキタスコンピューティングシステム、

ヒューマンコンピュータインタラクション、センサネットワーク等の研究に従事。ACM, 電子情報通信学会, IEEE 各会員。



河口 信夫 (正会員)

1990 年名古屋大学工学部電気電子工学科卒業。1995 年同大学大学院工学研究科情報工学専攻博士課程満了。同年同大学工学部助手、同大学講師、准教授を経て、2009 年より同大学大学院工学研究科教授。NPO 位置情報サー

ビス研究機構 Lisra 代表理事。モバイルコミュニケーション、ユビキタスコンピューティング、行動センシングの研究に従事。博士(工学)。ACM, IEEE, 人工知能学会, 日本ソフトウェア科学会, 電子情報通信学会, 日本音響学会各会員。本会シニア会員。