

# 研究終了報告書

## 「時系列信号の画像表現を用いた複雑行動認識」

研究期間：2020年11月～2023年3月

研究者：吉村直也

### 1. 研究のねらい

行動認識技術は、ユーザが装着したウェアラブルデバイスから得られる加速度などのセンサデータを解析することで、ユーザが行っている行動を推定する技術である。応用先の1つとして、スマート工場や物流センタにおける作業員のモニタリングがある。作業員のモニタリングを実現するためには、「ネジの取り付け」や「梱包資材の組み立て」といった「作業工程」の認識が求められる。しかし作業工程の認識には2つの課題がある。たとえば、「ラベルを貼る」といった1つの作業工程は、「ラベルを取り」、「貼り付け」、「擦って密着させる」という複数の動作から構成され、アイテムの個数などによって各動作の繰り返し数などが変化する。したがって、(1)複数の動作で構成され、継続時間にばらつきがある行動を認識可能な技術が必要である。本研究では、このような行動を「複雑行動」と定義する。しかし、工場における生産ラインでは作業員ごとに担当する作業が異なるだけでなく、ラインの見直しによって行動セットが変わる。そのため、(2)認識対象の行動セットが異なる大量のモデルの学習と更新を効率よく実行可能にすることも必要である。したがって、本研究のねらいは、作業工程をはじめとする「複雑行動」を少ない学習データで認識するための技術を確立することである。

本研究では、上記の目的を達成するため、1次元の時系列信号を2次元のテキスト画像に変換する「時系列信号の画像表現」に着目した。画像表現はセンサデータの繰り返しパターンを表現することに長けており、複雑行動を認識するために必要な、「動作」と「動作の組み合わせ」を同時に抽出するために非常に有用であると考えられる。本研究では、まず複雑行動を認識するベースライン手法を確立し、複雑行動の認識性能を検証することを目指す。そして、このモデルの構築を通して得た知見を活用しながら、時系列信号の画像表現を用いた複雑行動の認識技術の確立を目指す。

### 2. 研究成果

#### (1) 概要

ACT-X 研究の実施により、(1)複雑行動認識のベースラインモデルの開発、(2)時系列信号の画像表現を用いた複雑行動認識手法の開発、(3)複雑行動に注目した大規模マルチモーダル行動認識データセットの開発という3つの研究成果を得ることができた。

研究成果(1)について、限られた学習データで複雑行動である作業工程を認識することを目指す、「Lightweight Ordered-work Segmentation Network (LOS-Net)」を提案した。Dilated Convolution を活用した Work Process Context Pooling (WPCP) Module を提案し、作業工程を構成する動作とそれらの動作の組み合わせを効率的に抽出することができた。また、Boundary Detector によって行動の境界検出を行うことで、セグメンテーションの精度を高めた。さらに、行動の出現順序といった作業行動に関する事前知識を活用する Refinement Module を導入す

ることで、モデルの構築に必要なラベル付き学習データ量を削減した。

研究成果(2)について、まず (i)画像表現を用いたセグメンテーション手法の開発を行った。既存の画像表現を入力とするモデルはウィンドウごとの分類問題を想定しており、2次元の信号を入力し、出力は各クラスの確率を示すスカラ値である。一方で、複雑行動の認識にはセグメンテーションが必要であり、2次元の信号を入力し1次元の信号を出力することが求められる。そこで、ネットワーク上で2次元から1次元に変換しやすい入力信号の作成方法と、それを入力とする認識モデルを開発した。また、ImageNet など画像データを学習した事前学習済みモデルを特徴抽出器として複雑行動認識に活用するための研究もを行い、ResNet-18 の入力層に近いブロックの重みは行動認識にも適用可能であることがわかった。

研究成果(3)について、物流センターにおける梱包作業に着目した大規模マルチモーダルデータセット「OpenPack dataset」を構築した。16名の被験者の梱包作業を、加速度や深度画像を含む9種類のセンサで合計53時間分記録した。複雑行動を構成する基本動作に関するアノテーションを行い、複雑行動とその予測結果に関して詳細な分析が実現可能となった。また、OpenPack データセットは公開データセット<sup>1</sup>であり、その活用を促進するために PerCom 2023 のワークショップ BiRD<sup>2</sup>において、行動認識コンペティションを運営している(2023年1月15日まで)。

これら研究成果について、論文誌1件、国内研究会2件などの実績を得た。

[2-1-1] <https://open-pack.github.io/>

[2-1-2] <https://bio-navigation.jp/bird2023/>

## (2) 詳細

### ■■ 研究成果(1) 複雑認識のベースラインモデルの開発 ■■

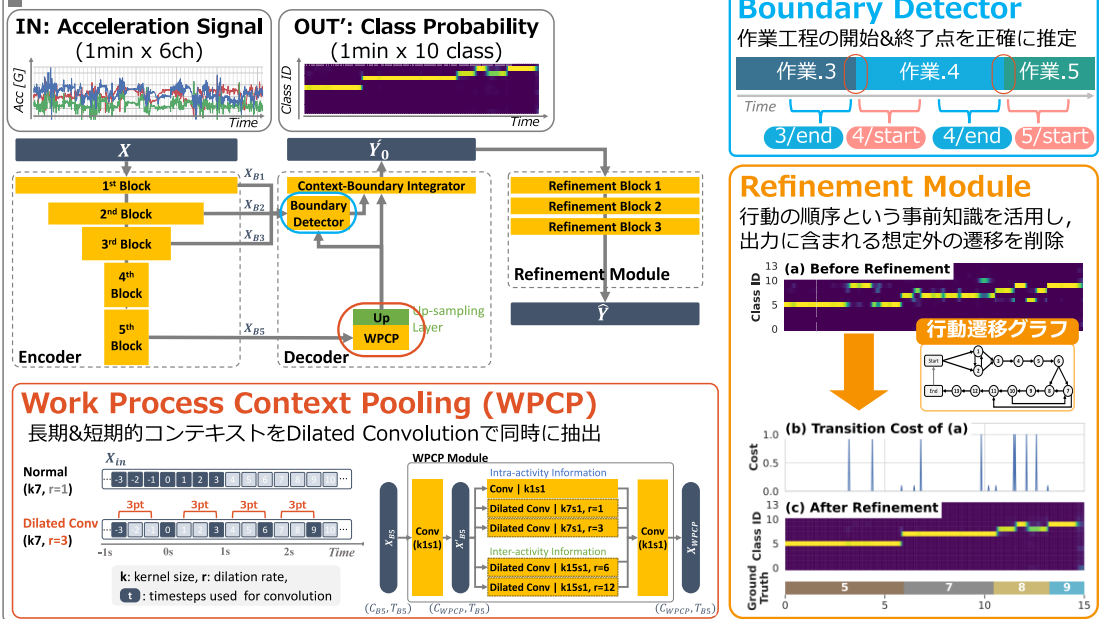
ウェアラブルセンサを用いて、工場や物流センターの作業員の作業行動を限られた学習データで認識する行動認識モデル Lightweight Ordered-work Segmentation Network (LOS-Net) を開発した(図1)。

認識対象の作業工程は複雑行動であり、例えば「箱の組み立て」工程は、「箱を開く」「フラップを折る」「ガムテープで止める」などの複数の動作で構成される。認識のためには、短期的コンテキストであるそれぞれの動作と、長期的コンテキストである動作の組み合わせに関する情報を抽出することが必要である。LOS-Net では Dilated Convolution を活用した Work Process Context Pooling Module を提案することにより、効率的に短・長期的コンテキストの抽出を実現した。この Dilated Convolution を用いるアイデアは、同時並行で進めていた研究成果(2)で得られた知見が活用されている。また LOS-Net では、限られた学習データ量でモデルを学習することに挑戦した。作業行動には行動の出現順序が原則決まっているなどの事前知識がある。LOS-Net では Refinement Module を導入することで事前知識を活用し、不足している学習データを補った。

実際の工場や物流センターで収集された11名の作業員のデータを用いて、提案手法を評価した。学習データが少ない環境下において、既存手法(U-Net)に対して、提案手法 LOS-Net は平均して10%程度F値を向上させた。

本研究成果は、ユビキタスコンピューティング分野におけるトップ国際論文誌である IMWUT に採択され(業績1-1)、国際会議 Ubicomp で口頭発表を行った。

図1: LOS-Net



研究成果(2) 時系列信号の画像表現を用いた認識手法の開発

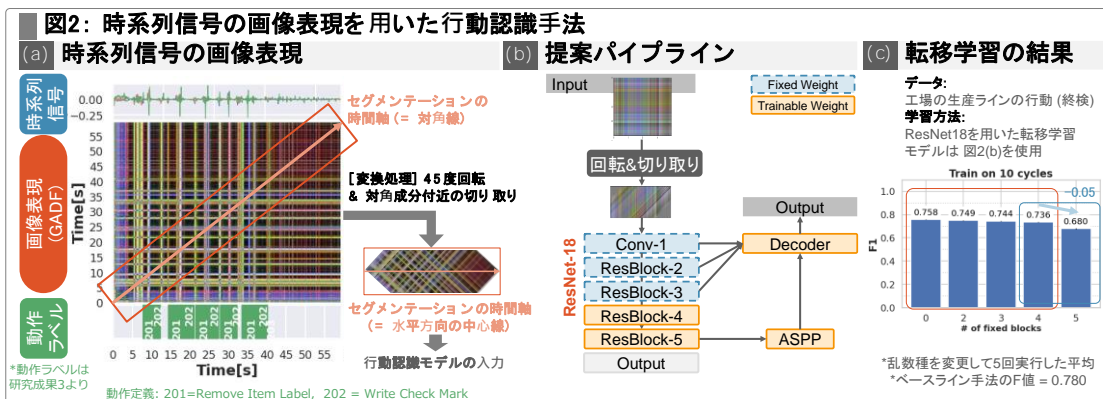
時系列信号の画像表現を入力とする複雑行動のセグメンテーション手法の開発と、提案モデルの学習手法に関する知見を得た。順に説明する。

時系列信号の画像表現は、1次元の時系列信号を図2(a)のように2次元に展開し、テキストチャ画像として扱う手法である。従来手法[Lu 2019]は、センサデータにスライディングを適用し、ウィンドウごとの認識を想定している。しかし、複雑行動の認識には、タイムステップごとに推定を行うセグメンテーションが必要であり、この問題に対し入力信号の加工法と、認識のパイプラインを提案した。時系列データのセグメンテーションでは、2次元のデータを1次元に変換する必要があるが、注目すべき時刻の要素が対角線上に分布しており、平均などを用いて1次元に圧縮すると、注目したい時刻以外の情報が多く含まれるため、短い行動の認識が非常に難しくなる。そこで、提案手法では、図2(a)のように入力画像を、対角成分を中心に切り取って再配置した。これにより、モデル内で長辺のサイズを保ったまま、短辺方向を圧縮するだけで、必要な過去と未来の情報を取り込みながら1次元のデータに変換し、注目した時刻の行動を推定することができる。この手法を物流センタで収集されたデータを用いて評価した。平均を使って2次元から1次元に変換するナイーブな手法では、F値が0.278であったのに対し、提案手法は0.320と、およそ0.04スコアを向上させた。

時系列信号の画像表現を用いる目的の一つは、動作をテキストチャとして扱うことで、動作に関する特徴抽出を、行動データではなく ImageNet などの一般の画像データから行うことにある。この目的の実現のため、一般の画像データを学習したモデルの転移学習に関する調査を行った。提案手法の Encoder 部分 (図2(b)) に学習済み ResNet-18 を使用し、fine-tuning するブロックを変更しながら作業行動データを学習した。結果を図2(c)に示す。入力層に近いブロック (Block1~3) では、ベースライン手法に近い F 値となった。一方で、Block4 と Block5 の間に大きな認識精度の低下が見られた。この実験より入力に近い部分であれば、一般の画像か

ら抽出した特徴を行動認識に活用できることが分かった。

一方で、研究を進める上で既存のデータセットの課題が見えてきた。本課題を進める際の大きな障害となりうる、「動作に関するアノテーションが無いためテキストと動作の関連を調べることができない」、「被験者数が少なく、テキストの被験者間比較ができない」などの問題を解決するため、研究成果(3)となるデータセットの構築を行った。

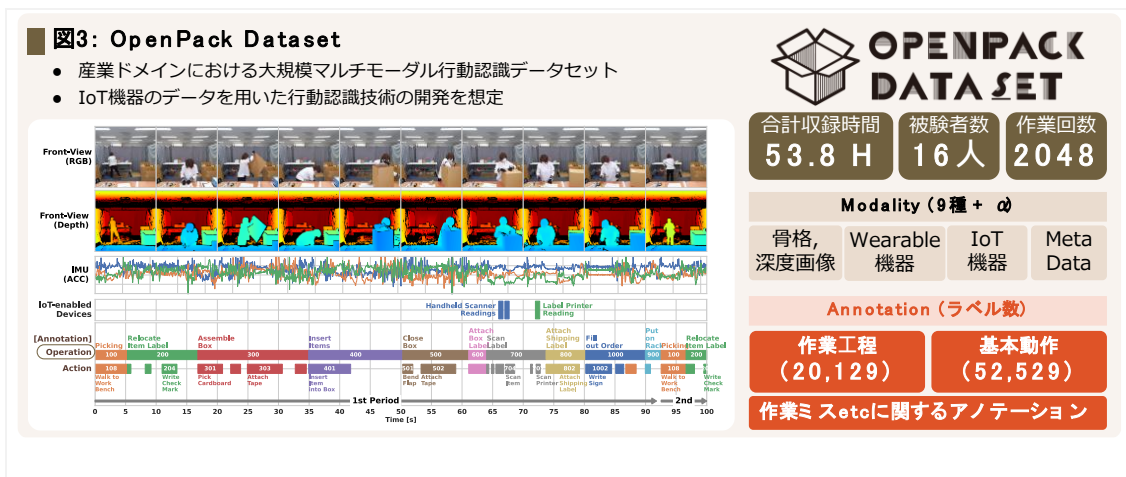


### 研究成果(3) 大規模な複雑行動データセットの開発

物流センタにおける梱包作業に注目した、大規模マルチモーダルデータセット「OpenPack Dataset」を構築した (図 3, 業績 3-1)。16 名の被験者から、加速度や深度画像を含む 9 種類のセンサで梱包作業中の行動データを収集した。合計収録時間は 53 時間、収集した作業回数 は 2048 回分である。また、OpenPack はより近代的な工場や物流センタを想定しており、スキャナをはじめとする IoT 機器のデータや、オンライン注文管理システムを想定した注文票 (各注文で梱包するアイテムのリスト)、梱包するアイテムのサイズや JAN コードなどのメタデータを多数提供する。作業行動のデータセットとして LARa [Niemann 2022] が公開されているが、合計収録時間は 15 時間程度、作業の繰り返し数は 499 回で、RGB 画像・骨格データなど合計 4 種類が収録されている。収集したデータ量とセンサの種類観点においても、OpenPack が既存データセットを大きく上回っていることがわかる。

行動クラスは、「Picking, Assemble Box」などの作業工程を 10 種類、作業工程を構成する動作 32 種類を定義しアノテーションを行った。ラベル数は作業工程が 20,129 件、基本動作が 52,529 件である。これに加え、研究課題(1)(2) で課題として見えてきた、作業ミスや作業と関係の無い動作に関するアノテーションも整備している (2022 年 12 月時点で半分完了)。また、この動作ラベルを、使用することで研究課題(2)の時系列信号の画像表現におけるテキストと動作の関連性を検証することができた。

OpenPack データセットの活用を推進するために PerCom2023 のワークショップ BiRD において、OpenPack データセットを用いた作業行動の認識コンペティション「OpenPack Challenge 2022」を運営している。本コンペは 2022 年 10 月 15 日開始、2023 年 1 月 15 日終了の予定である。2022 年 12 月 19 日の時点で、予測結果の投稿サイト Codalab 上での参加登録者数は 58 名である。アカデミアだけでなく、産業界からの参加もあり、国内外も含めて幅広い層に ACT-X 研究の成果をリーチすることができている。



### 3. 今後の展開

ACT-X 研究の目標は作業行動の認識技術の確立であり、その基盤技術は研究成果(1)・研究成果(2)で確立できたと言える。これらの研究で明らかになった課題の解決を目指して、研究成果(3)の OpenPack データセットの構築を行い、様々な課題の提起を行なった。今後は、他の研究者とも協力しながら、5 年程度のスパンで、(i)注文などのメタデータの活用や、(ii)作業ミスや想定外動作の検出といった課題に取り組んでいきたいと考えている。

また、作業行動の認識技術の実際の工場での運用を目指す活動にも力を入れることが必要であると考えている。こちらは 10 年程度の長期的なスパンで、研究者だけでなく、産業エンジニアの人々と連携をとり、OpenPack データセットを軸にしながら、必要なデータの提供や技術開発、実用化のための新たなタスクの定義などに取り組んでいきたい。

### 4. 自己評価

#### ■■ (1) 研究目的の達成状況 ■■

ACT-X の研究目的は、「複雑行動を少ない学習データで認識する技術の開発」であり、その手段として「時系列信号の画像表現を用いた行動認識技術の確立」を設定した。主目的の複雑行動の認識は、研究成果(1)の通り、国際論文誌IMWUTにも採択されたことから、十分に達成できたと言える。また、時系列信号の画像表現を用いた行動認識技術は、ベースラインに近いパフォーマンスに到達できた。ただし、研究成果(1)で提案した手法のパフォーマンスが非常に高く、それを超えることは難しかった。しかし、時系列信号の画像表現という難しい問題から得た多くの知見が、研究成果(1)で提案したモデルに数多くの貢献をした。また、研究成果(1)(2)で明らかになった問題点の解決に向け、研究成果(3)のデータセットの構築という、追加の課題に取り組み、行動認識コンペティションの運営というアウトリーチ活動も実施することができた。

研究費に関しては、研究を遂行する上で過不足なく活用することができた。特に、データセットの構築の際は、ACT-X の研究費が非常に大きなサポートとなった。

#### ■■ (2) 研究体制 ■■

2020 年 10 月から 2023 年 3 月の全期間、大阪大学にて研究を実施した。また、年 2 回の領域会議、および 年度末のサイトビジットにより、アドバイザーである東京大学川原教授を始め、アドバイザーの先生の皆様や他の ACT-X 研究者にアドバイスをいただき、研究を進めた。研究成果

(3)に関しては、ACT-X 2期生の大川氏と個別ミーティングを持ち、情報交換を行った。継続的に同年代の研究者から多くの刺激を受けることができ、非常に貴重な体験となった。

### ■■ (3) 研究成果の科学技術および社会・経済への波及効果 ■■

研究成果(1)については、工場や物流センタにおける作業行動の認識を可能とする手法を提案することで、ウェアラブルデバイスを用いた手法でも作業員のモニタリングが可能であることを示すことができた。また、研究成果 3 のデータセットの構築と行動認識コンペティションの運営によって、産業エンジニアが作業行動の認識技術に気軽に触れ、公開データセットで独自のモデルを評価することができる環境を整備することができた。行動認識コンペティションには、アカデミアだけでなく、多くの企業の方が参加・興味を示してくださっており、今後、作業行動の認識技術の現場への導入の検討が加速することが期待される。

また、OpenPack データセットは作業行動の認識だけでなく、研究成果(1)(2)を通して得られた問題点や知見が反映されている。作業ミスの検出や、想定外の動作への対応など、新たなタスクの提案と検証ができるようになっており、重要な学術的な貢献だと考えられる。

## 5. 主な研究成果リスト

### (1) 代表的な論文(原著論文)発表

研究期間累積件数:1件

1. Naoya Yoshimura, Takuya Maekawa, Takahiro Hara, Atsushi Wada, Yasuo Namioka, “Acceleration-based Activity Recognition of Repetitive Works with Lightweight Ordered-work Segmentation Network,” Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies, 2022, vol.6, no.2

工場や物流センタの作業員を想定し、ウェアラブルデバイスを用いて、限られた学習データで作業員の作業行動を認識するモデル Lightweight Ordered-work Segmentation Network (LOS-Net) を提案した。LOS-Net は、Dilated Convolution を活用した短期的・長期的コンテキストを効率的に抽出するモジュールによって、複数動作で構成される作業工程の認識性能を大きく向上させた。また、行動の出現順序などの事前知識を活用するモジュールを導入することで、モデルの構築に必要な学習データ量を削減した。

### (2) 特許出願

研究期間全出願件数:0件(特許公開前のもも含む)

### (3) その他の成果(主要な学会発表、受賞、著作物、プレスリリース等)

1. 吉村 直也, Jaime Morales, 前川 卓也, 原 隆浩, “OpenPack データセット: 物流センタにおける 大規模マルチモーダル行動認識データセットの構築,” 第 193 回 DPS 研究会 研究報告, 2022 年 12 月 (口頭発表)
2. 吉村 直也, 前川 卓也, 原 隆浩, 和田 篤, 浪岡 保男, “国際発表奨励賞,” 第 76 回 UBI 研究会, 2022 年 11 月, 兵庫. (受賞・ポスター発表)
3. Naoya Yoshimura, Takuya Maekawa, Takahiro Hara, Atsushi Wada, Yasuo Namioka, “Acceleration-based Activity Recognition of Repetitive Works with Lightweight



Ordered-work Segmentation Network,” Ubicomp, Sep. 2022, UK. (口頭発表)

4. 吉村 直也, 前川 卓也, 原 隆浩, 和田 篤, 浪岡 保男, “工場・物流センター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,” 第 74 回 UBI 研究会, 2022 年 6 月, 福岡. (口頭発表)
5. 吉村 直也, 前川 卓也, 原 隆浩, 和田 篤, 浪岡 保男, “優秀論文賞,” 第 74 回 UBI 研究会, 2022 年 6 月, 福岡. (受賞)