

宅内行動アノテーションの省力化のための行動認識手法の提案と評価

浅井俊宏^{†1}

奈良先端科学技術大学院大学

松井智一^{†2}

奈良先端科学技術大学院大学

諏訪博彦^{†3}

奈良先端科学技術大学院大学

安本慶一^{†4}

奈良先端科学技術大学院大学

1. はじめに

高齢化が進む現代社会において、家庭内における高齢者の生活状況を把握し、ヘルスケアや見守り支援を行うための技術が注目されている。具体的には、家庭内での移動、活動時間、睡眠パターンなどをモニタリングすることにより、健康状態の変化や異常を早期に検知し、迅速な対応や適切な介護計画の立案が可能となる。また、これにより、高齢者の生活の質 (Quality of Life: QOL) を向上させることが期待されている。このような行動認識を実現するためには、まず高精度なモデルを構築する必要があるが、その過程においていくつかの課題が存在する。

このような行動認識を実現するためには、まず高精度なモデルを構築する必要があるが、その過程においていくつかの課題が存在する。第一に、モデルの学習に必要な大量のラベル付きデータの収集が極めて困難である点が挙げられる。従来の研究では、被験者が日常的に行う行動に対して詳細なアノテーションを付与する必要があったが、この作業には多大な時間と労力がかかる。また、被験者自身がアノテーションを行う場合、その正確性にも課題が残る。第二に、家庭環境の多様性が、行動認識モデルの精度向上を阻む要因となっている。家庭ごとに生活様式が異なるため、汎用的なモデルでは各家庭の特有のパターンを適切に捉えることが難しい。一方、家庭ごとに個別のモデルを構築しようとすると、都度アノテーション作業が必要となり、実用化が遠く。そのため、アノテーションコストを最小限に抑えつつ、高精度なモデルを効率的に構築する手法が求められている。

さらに、行動認識技術は、非接触型であることやプライバシー保護の観点からも考慮する必要がある。近年、環境センサや人感センサを利用した非接触型のセンシング技術が注目を集めており、これらを活用することで、プライバシーに配慮したデータ収集が可能となる。このような技術の発

展は、ラベル無しデータを活用した新しい学習手法の適用可能性を広げるものである。

そこで本研究では、家庭内に設置した非接触型センサのセンシングデータを活用し、少量の行動ラベルでも効率的にモデルを学習できる手法を提案する。具体的には、無ラベルデータに対する自己教師あり学習 (Self-Supervised Learning) と、分類器の不確実性に応じた追加アノテーションを行う能動学習 (Active Learning) を組み合わせる。これにより、(1) 十分量のデータがあってもラベルが不足している問題への対処、(2) プライバシーに配慮しつつ行動認識の性能を向上、(3) 被験者のアノテーション負担軽減といった効果が期待できる。

本研究では、複数世帯 (4 家庭) を対象として、家庭内センサデータを収集し、そのうち 1 世帯分のデータに対して自己教師あり学習で無ラベルデータから分類モデルを実装した。さらに、能動学習として、分類モデルの確信度が低いサンプルのみ使用した学習を追加で行う事で、効率的に分類モデルの精度が向上されることを確認した。本論文では、関連研究を概観したうえで、提案手法の詳細と実験的評価を報告し、今後の課題を考察する。

2. 関連研究

2.1. 宅内行動認識

近年、行動認識 (Human Activity Recognition, HAR) 技術の発展により、スマートホーム、ヘルスケア、異常行動検知などの幅広い応用が期待され、高齢者の見守り等への応用が注目されている。[1, 2, 3, 4] これらの HAR を行うシステムは、ウェアラブルデバイスを利用した手法と利用しない手法の二つに大別することができる。

ウェアラブルデバイスを利用した行動認識は、加速度センサやジャイロスコップを搭載したデバイスを装着し、利用者の動きをセンシングすることで行動を認識する手法である。この手法は、デバイスの装着によって個人ごとの行動識別が可能であり、特に転倒検知や健康モニタリングの分野で多くの研究が行われている。Giulia らは、ウェアラ

In-Home Activity Recognition for Reducing Annotation Costs

^{†1} Toshihiro Asai, Nara Institute of Science and Technology

^{†2} Tomokazu Matsui, Nara Institute of Science and Technology

^{†3} Hirohiko Suwa, Nara Institute of Science and Technology

^{†4} Keiichi Yasumoto, Nara Institute of Science and Technology

ブルデバイスを活用し、状況認識を組み合わせた HAR システムを提案した [5]。しかし、装着の煩雑さや長期間の利用におけるデバイスの電池寿命、ユーザの装着拒否といった課題がある。

ウェアラブルデバイスを利用しない行動認識の手法は、近年、特にプライバシー保護の観点から注目されている。Xiaolong らは市販の Wi-Fi デバイスを活用した非接触型の人間の行動認識システムを開発し、LSTM ベースの時系列データ解析に、周波数情報を組み合わせた新しい認識モデルを提案した [6]。このような電波を活用した手法は、非装着型の利便性が高い一方で、センシング環境に依存しやすいという課題がある。例えば、家具の配置変更や機器の追加・削除が発生した場合、再キャリブレーションが必要になることが多い。

2.2. 人感・環境センサによる宅内行動認識システム

設置型の人感・環境センサを利用した行動認識は、非装着型の利便性が高く、ユーザの負担を軽減する手法として注目されている。ただし、Wi-Fi や RFID を使用する手法と異なり、住環境で収集されるデータは非常にプライベートなデータであることから、プライバシーやセキュリティに配慮する必要がある。

プライバシーに配慮しつつ、一般家庭における行動認識を目的としたセンシングシステムとして、CASAS のスマートホームキットがある [7]。CASAS は、居住者の動作や生活パターンを非接触かつ間接的に記録するプライバシーウェアなセンシングキットである。しかしながら、CASAS にはいくつかの課題も存在する。すべてのセンサはバッテリー駆動であり、定期的なメンテナンスが必要となるほか、バッテリー搭載によりセンサのサイズがやや大型化している。さらに、行動ラベルのアノテーションには、居住者自身が PC 上のソフトウェアを用いて行動ラベルを設定する必要がある、アノテーション作業の手間が発生する。

CASAS の課題を解決するため、同様のスマートホームキットとして SALON [8] が提案されている。SALON は、エネルギーハーベスティング（環境発電）技術を活用したセンサを採用し、長期間のメンテナンス不要な動作を実現している。これにより、バッテリー交換の必要がない。さらに、SALON は居住者の負担を軽減するため、物理的なアノテーションボタンを導入し、ユーザが直感的に自身の行動を記録できるようになっている。このアノテーションボタンも環境発電技術を利用しており、電池交換なしで動作するため、CASAS と比較してさらなるアノテーション負荷の軽減およびメンテナンスフリーの利点を有する。しかしながら、依然としてアノテーション作業は負担のかかる作業であり、実際に SALON ではアノテーション忘れによるラベルの欠

損が課題として残っている。

2.3. アノテーションの効率化・省力化

SALON ではアノテーションボタンを導入することで一定の負担軽減を実現したが、それでもすべての行動に対して正確なラベルを付与することは依然として難しく、さらなる省力化の工夫が求められる。この課題を解決する手法の一つとして、不確実性サンプリング（Uncertainty Sampling）がある [9]。

不確実性サンプリングでは、モデルが予測に自信を持ってないデータ点を優先的に選択し、必要最小限のアノテーションでモデルの精度を向上させる。例えば、行動認識において、モーションセンサの発火パターンが曖昧な場合や、異なる行動の境界が不明瞭な場合に、そのデータのみを居住者に確認してもらい、ラベル付けを行うことで、不要なアノテーション作業を削減しつつ、効率的に学習データを強化することが可能となる。

この不確実性サンプリングを学習に応用した手法として能動学習（Active Learning）が注目されている。能動学習は、機械学習モデルが高い確信度を持たないデータに対して、追加のアノテーションを優先的に要求する手法であり、アノテーション作業を効率化することができる [10]。

2.4. 本研究の位置づけ

従来の宅内行動認識システムでは、ウェアラブルデバイス、カメラ、マイクを活用する手法が広く研究されてきた。しかしながら、ウェアラブルデバイスは装着の煩雑さや電池寿命の制約といった課題があり、カメラやマイクを用いた手法はプライバシー侵害の懸念が大きい。一般家庭への導入が難しい。これらの課題を解決するため、非接触かつプライバシーに配慮したセンサを活用した手法が近年注目されている。

しかしながら、依然としてアノテーション作業の負担が大きいという課題が残る。特に、長期間のデータ収集を行う実環境においては、居住者が自身の行動を逐一記録する負担が大きく、アノテーション忘れやラベルの欠損が生じやすい。

本研究では、これらの課題を踏まえ、SALON でも使用したセンサを活用したプライバシー侵害の少ない宅内行動認識システムを構築するとともに、アノテーション作業の負担を軽減する新たな手法を提案する。具体的にはラベル無しデータを活用した自己教師あり学習と能動学習を組み合わせることで、少量のアノテーションデータでも高精度な行動認識が可能なモデルを構築する。また、実環境における長期間のセンシングデータを用いた評価を行い、提案手法の有効性を検証する。これにより、実環境でも宅内行動

認識の普及を妨げる要因の一つであるアノテーションの負担を軽減し、より実用的なシステムの実現を目指す。

3. 提案システム

本システムのセンシングについては、ユーザの日常生活に干渉せずに行動データを収集することが求められる。具体的には、環境センサや人感センサを用い、家庭内に設置されたセンサ群から情報を取得する仕組みを採用する。これにより、身体に装着するデバイスを用いる必要がなく、ユーザが意識することなく行動記録を行うことが可能となる。また、データ収集において、ユーザのプライバシーを保護することは極めて重要である。特に、映像や音声など個人の特定が可能な情報の使用を避けるため、温度、湿度、移動感知等の個人特定ができないデータを用いる。

3.1. システム構成

提案システムの要件を満たすよう、使用したデバイスを図 1 に示す。

3.2. 人感センサ

人感センサは、赤外線の入射変化を検知することで、人間等の熱源の動きを感知する。メンテナンスフリーかつ長期間の運用を可能とするため、低消費電力通信規格である EnOcean に対応した ROHM 社の温度センサ STM431J と Panasonic 社の人感センサモジュール EKMB1101112 を組み合わせて使用した。このセンサは、環境発電のみで動作し、データ収集および送信が可能である。

加えて、環境発電が困難な状況においても安定したデータ取得を実現するため、予備の人感センサとして Switchbot 製を採用した。BLE 通信を利用し、単 4 電池 2 本で動作するこのセンサは、通常の使用条件下で最大 3 年間の稼働が可能である。

3.3. 環境センサ

環境センサとして、小型で多機能な Omron 社の 2JCIE-BL01 (BAG タイプ) を使用した。このセンサは、BLE 通信を利用して温度、湿度、照度、気圧、騒音などのデータを記録する。また、UV index や熱中症警戒度といった補助的な情報も取得可能である。1 つのボタン電池 (CR2032) で動作し、サンプリング周期を 5 分に設定した場合、電池寿命は約 6 か月である。

3.4. ドアセンサ

ドアセンサは、ROHM 社製の磁気ドア開閉センサ STM250J を採用した。このセンサは、ドアとドア枠に設置



図 1 システム構成

されたセンサ本体とマグネットの位置関係を検知し、ドアの開閉状態を記録する。EnOcean 規格に対応しており、環境発電による動作が可能である。

3.5. データサーバ

データサーバには、小型 PC である Intel 社の NUC を利用した。このサーバには、BLE 通信用ドングルおよび EnOcean 通信用ドングルが接続され、各センサがブロードキャストするデータを収集する。データは MongoDB に格納され、事前に登録された EEP や BLE MAC アドレスに基づいてセンサ個体情報を識別し保存する。また、ピクセラ社製の PIX-MT100 (LTE 対応 USB ドングル) 等を利用してオンライン接続を可能とし、無線 LAN 環境が整っていない場所でもサーバ監視やデータダウンロードが行える。

3.6. 自己教師あり学習

本研究では、無ラベルデータから特徴量を抽出するために、自己教師あり学習の一手法である DINO (Distillation with No Labels) [11] を採用する。自己教師あり学習は、データに明示的なラベルを付与することなく特徴表現を獲得する手法であり、代表的な手法として SimCLR[12] や MoCo[13] などのコントラスト学習に基づく手法が挙げられる。これらの手法はポジティブサンプルとネガティブサンプルを定義する必要があるが、宅内行動認識のようなタスクにおいて、異なる時間帯や異なるセンサのデータがどの程度類似しているかを明確に定義することは難しい。DINO はポジティブ・ネガティブの関係を明示的に設定する必要がなく、時系列の家庭内行動データから特徴量を取得する本研究のタスクに適していると考えられる。

DINO の特徴量抽出においては、教師ネットワークと生

徒ネットワークの2つのニューラルネットワークを用いる。学習の際、同一の入力データに対し異なるデータオーグメンテーションを適用し、教師モデルの出力する特徴量に生徒モデルの出力を近づけるように学習が進む。この過程において、教師ネットワークのパラメータは勾配更新を受けず、蒸留損失を最小化する形で更新される。この仕組みにより、DINO は自己蒸留による特徴学習を行い、無ラベルデータから有用な特徴量を獲得することができる。

今回の DINO の学習では、以下の手順を実施する。

- 収集した非接触センサデータを一定の時間スライスにリサンプリングする
- データオーグメンテーションを適用し、2つの異なるビューを生成する
- それぞれのビューを、教師ネットワークと生徒ネットワークに入力する
- 教師ネットワークの出力を基準として、生徒ネットワークの出力がこれに近づくように損失を計算し、学習を行う

このアプローチにより、DINO で事前は無ラベルデータを用いた特徴表現の学習を行うことで、家庭の時系列的な特徴量を取得する。

3.7. 能動学習

本研究では、少量のアノテーションデータを効果的に活用しながら高精度な行動認識を実現するために、能動学習を用いた分類器の更新手法を導入する。能動学習は、機械学習モデルの不確実性が高いデータを選択的にアノテーションすることで、限られたラベル付きデータの中で最大限の性能向上を図る手法である。本研究では、DINO によって事前学習された特徴量を用いて分類器を構築し、分類器の確信度をもとに追加アノテーションを行うデータを選択することで、アノテーションの省力化を実現する。

本研究では、DINO で抽出された特徴量を用いて分類器の初期学習を行う。この段階では、すべてのラベル付きデータを用いるのではなく、限定的なラベル付きデータのみを利用する。その後、初期学習した分類器を用いて、さらなる学習を行うべきデータの選定を行う。具体的には、分類器の出力する確信度のスコアを算出し、確信度が一定の閾値を下回るデータを不確実なデータとして抽出する。本研究では、Softmax による確率値の最大値を基準とし、分類器が明確な判断を下せないデータを対象として選択する。このように、モデルの判断が曖昧なデータに対して追加アノテーションを行うことで、分類器の性能を効率的に向上させることを目指す。

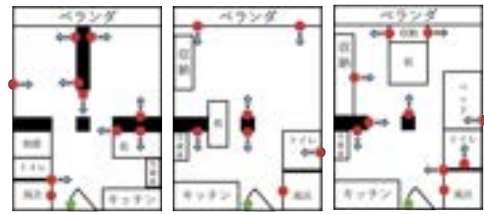


図2 2人居住者のセンサ設置箇所



図3 単身居住者のセンサ設置箇所



図4 人感センサおよび環境センサの設置例



図5 ドアセンサの設置例

4. 実験

4.1. 実験目的および手順

本実験の目的は、提案システムが一般家庭において想定通り運用可能であることを確認し、得られた実環境のデータに対しても、提案手法を用いることでアノテーションコストの省力化が可能であることを検証する事である。

実験対象は、大阪府東大阪市宝持西の集合住宅に在住している家庭4件であり、うち3家庭が2人居住者、1家庭が単身居住者である。

4.2. センサ設置要件

各家庭のセンサの設置箇所について図2、図3に示す。なお、センサ設置箇所は赤丸で、矢印でセンサの向きを、ドアセンサは緑丸で表現している。それぞれのセンサの設置例を、図4、図5に示す。

環境発電で動作する人感センサは、各家庭に配置する数を制限するために、10 個程度用意し任意の個所に設置した。ただし、設置個所はベランダ・寝室・居間・キッチン・トイレ・浴室など居住者が頻繁に使用する個所を優先的に設置した。なお、耐水ではないため、浴室に設置する際は浴室前のドアに、地面を見下ろす形で設置した。また Switchbot 製人感センサは 5 個程度用意し、就寝時の寝室など、環境発電が行えないような状況でも利用する部屋に対して優先的に設置した。

環境センサは、モーションセンサと同様の方法で、基本的には、モーションセンサの隣に配置することとしている。サンプリング周期は、3 分間に一度と設定した。なお、設置個所の利用頻度が高くなく、他の環境センサとの距離が近い場合、環境センサは省略し、人感センサのみ設置する事としている。

ドアセンサは、人感センサと同様の方法で玄関に設置した。この際に、ドアの枠側に開閉を検知するためのマグネットを設置し、ドアの可動部側に本体を設置した。

データサーバは、通信強度を高く保つために、家庭内のなるべく中心に近い箇所に配置した。その際、机の下、電話機の横など、邪魔にならない場所へ配置した。

5. データ収集および解析の結果

2024 年 4 月以降にデータ収集に同意した被験者に対して、システムの設置を行った。その上で、実験として単身家庭の 1 週間の ADL データを使用して、DINO の学習と能動学習を組み合わせた行動認識を行った。取得したデータの例として、単身居住者の一日のセンシングデータを図 7、図 8、図 9、図 10 に示す。なお、人感センサおよびドアセンサの反応回数については、10 秒間の反応回数をプロットした。

なお、本研究ではアノテーションの省力化を検証するため、DINO によって抽出された特徴量を用いた分類器の学習を実施した。DINO の学習には無ラベルデータのみを使用しているため、分類器の学習および評価にはラベル付きデータが必要となる。そこで、本研究では被験者の負担を考慮し、行動ラベルの付与を第三者によって行った。

アノテーションの手法としては、時系列センサデータに対し、特定の行動パターンが明確に識別できる時間帯を選定し、適切な行動ラベルを付与する形を取った。具体的に、以下の基準に従い、アノテーションを行った。

- トイレ内の EnOcean 人感センサの反応が顕著である時間帯には、照明を点灯して使用している事が想定されるため「トイレ」のラベルを付与した（プロットの背景色が赤色の期間）



図 6 t-SNE による 4 日間の特徴量の可視化

- 朝と昼のキッチン周辺的环境センサの騒音上昇と、人感センサの反応が顕著である時間帯には、料理を行っている事が想定されるため「料理」のラベルを付与した（プロットの背景色が緑色の期間）
- 夜間の長時間にわたる静的な環境変化、特に人感センサの反応が隠微である時間帯には、就眠したことが想定されるため「睡眠」のラベルを付与した（プロットの背景色が青色の期間）
- それ以外の時間帯には「その他」のラベルを付与した

次に、4 日間の ADL データを DINO で学習した。データの前処理として、環境センサの温度・湿度・照度・気圧・騒音のデータを 10 秒ごとにリサンプリングし、異常値の除去および欠損値を線形補完した。人感センサ・ドアセンサは 10 秒ごとの反応回数をカウントし、これを環境センサのデータと統合し、入力する時系列特徴量とした。その後、3 章で提案した手法に則し、DINO の学習を行った。生成された特徴量の有用性を検証するため、t-SNE を用いて次元削減を行い、特徴空間を 2 次元で可視化した。可視化した図を図 6 に示す。なお、それぞれのプロットは、図 7、図 8、図 9、図 10 と同じく、赤色がトイレ、緑色が料理、青色が睡眠を示す。

最後に行動予測を行うため、作成した DINO モデルの次層に全結合層を追加し、分類器としての訓練を行った。訓練に使用したデータは能動学習を再現するため、以下の 3 パターンに分割して行った。

手法 1: ベースライン

最小限のラベルデータを使用するパターンであり、1 日目と 2 日目のデータを用いて分類器を学習し、5 日目から 7 日目のデータを評価に用いた。この手法では、最低限のラベルデータのみを使用し、どの程度の精度が得られるかを確認した。

手法 2: 完全なアノテーション

全ラベルデータを使用するパターンであり、1 日目から 4 日目のすべてのラベル付きデータを用いて分類器を学習し、5 日目から 7 日目のデータで評価を行った。この手法は、ラ

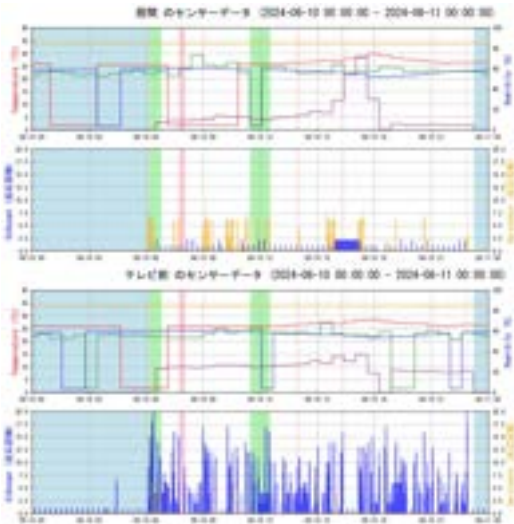


図7 単身居住者の1日のセンシング結果
(居間・テレビ前)

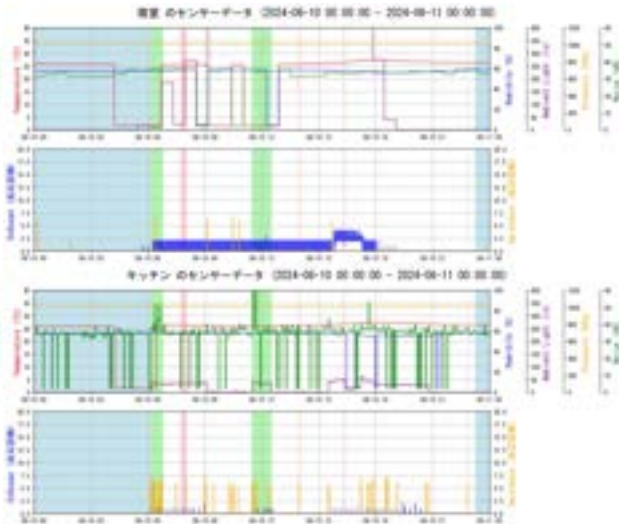


図8 単身居住者の1日のセンシング結果
(寝室・キッチン)

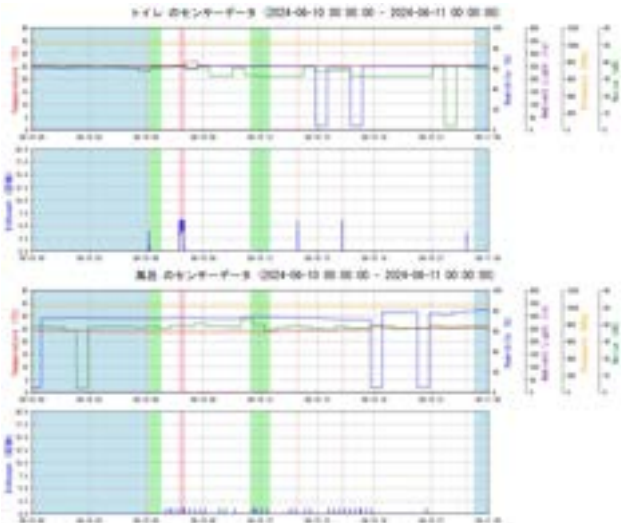


図9 単身居住者の1日のセンシング結果
(トイレ・風呂)

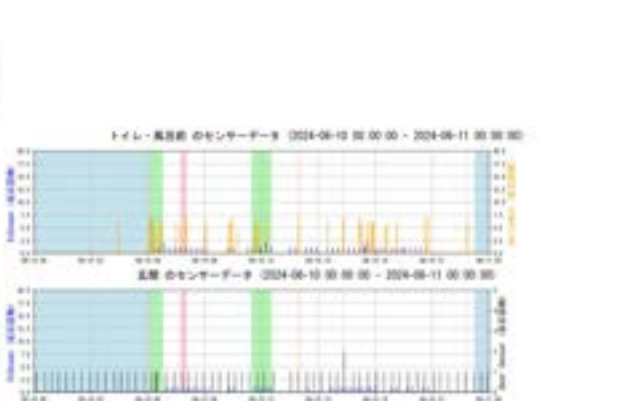


図10 単身居住者の1日のセンシング結果
(トイレ・風呂前および玄関)

ベル付けを最大限に活用した場合の上限性能を測定するための比較対象とした。

手法3: 能動学習

手法1に能動学習を適用するパターンであり、1日目と2日目のデータで分類器を学習した後、分類の確信度が低い3日目と4日目のデータを選択的に学習に使用した。この手法により、最小限のラベル付けで分類精度を向上させられるかを検証した。それぞれの手法の結果を表1に示す。また、図11、図12、図13にそれぞれの手法で得られた混同行列を示す。なお、図11、図12、図13の行動ラベルは、それぞれ1:睡眠、2:トイレ、3:料理、4:その他に対応して

いる。

表1 各手法の行動推定結果

	Accuracy	Avg-Recall	Avg-Precision	Avg-F1-score
手法1	0.902	0.902	0.895	0.890
手法2	0.931	0.931	0.931	0.926
手法3	0.909	0.909	0.904	0.900

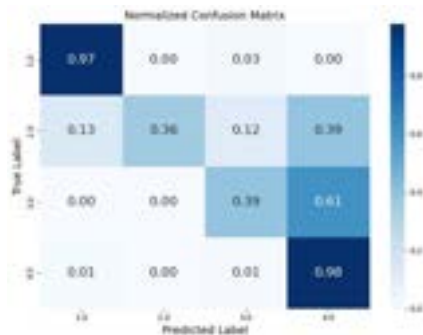


図 11 手法 1 の混同行列



図 12 手法 2 の混同行列



図 13 手法 3 の混同行列

6. 考察

取得されたデータについては、居間・テレビ前・寝室はベランダ側にあり、日中は頻繁に反応する結果となった。さらに、光の弱まる夜間については、EnOcean 製の人感センサの反応回数が減少する結果となった。テレビ前の人感センサについては、環境発電ができず、被験者が就寝していると思われる時間帯でも一定間隔で反応している。これは EnOcean 製の人感センサの元となった温度センサの特性で、蓄電池を搭載しているため、日中に取得できた余剰のエネルギーで動作していると考えられる。

一方で、キッチン・トイレ・風呂などの個所のセンサは、集合住宅における廊下側に位置している。そのため、ベラ

ンダ側よりエネルギーがあまり取得できていないため、反応回数がベランダ側のセンサと比較して少ないと考えられる。特に、トイレは使用していない際は消灯し、使用する際のみ点灯するため、センサ反応がある時、ない時がはっきりと分かれる結果となったと考えられる。

キッチンおよびトイレ・風呂前の EnOcean 製人感センサも反応が少ない結果となったが、ここでは Switchbot 製の人感センサが頻繁に反応している。電池駆動しているセンサを使用したため、環境に左右されずに反応することができたと考えられる。

環境センサのデータについては、温度が数度まで低下する期間が存在するなど、全体的に不連続な点が多い結果となった。これは、Omron 社の環境センサのデータは、測定データを一連の 16 進数の文字列として送信するため、何らかの原因で文字列が欠損した結果であると考えられる。

行動予測については、表 1 より、手法 3 の全体の性能が手法 1 より向上しており、能動学習の効果が表れたと考えられる。ただし、図 11、図 12、図 13 よりすべての手法において、「睡眠」と「その他」の行動は予測できているのに対して、「トイレ」、「料理」の行動は精度が低い。これについては、図 ?? から、「睡眠」や「その他」の行動がほとんどを占めており、逆に「トイレ」、「料理」等の行動が少ない事に起因していると考えられる。

7. おわりに

本研究では、家庭内の行動認識におけるラベル付け作業の負担を軽減するため、自己教師あり学習と能動学習を組み合わせた行動認識手法を提案した。具体的には、人感センサや環境センサから得られる無ラベルデータに対して、DINO を用いた自己教師あり学習を行い、少量のラベル付きデータと能動学習を組み合わせることで、高精度な行動認識モデルを構築することを目指した。実際の一般家庭にセンサを設置し、長期間にわたるセンシングデータを用いた評価実験では、提案手法がラベルデータを最小限に抑えながらも、従来手法と同等以上の分類精度を達成できる可能性を示した。特に、分類器が確信度の低いデータのみをアノテーション対象とする能動学習により、被験者や研究者のラベル作業コストを抑制できる点が確認された。今後は、複数居住者を含むケースへの適用や、より多彩な行動カテゴリへの拡張、さらには学習アルゴリズムの改良による行動クラスタの分離精度向上を検討していく必要がある。

参考文献

- [1] Wang, J., Chen, Y., Hao, S., Peng, X. and Hu, L.: Deep learning for sensor-based activity recognition: A survey, *Pattern Recognition Letters*, Vol. 119, pp. 3–11 (online), <https://doi.org/10.1016/j.patrec.2018.02.010> (2019). Deep Learning for Pattern Recognition.
- [2] Gu, F., Chung, M.-H., Chignell, M., Valaee, S., Zhou, B. and Liu, X.: A Survey on Deep Learning for Human Activity Recognition, *ACM Comput. Surv.*, Vol. 54, No. 8 (online), 10.1145/3472290 (2021).
- [3] Sun, Z., Ke, Q., Rahmani, H., Bennamoun, M., Wang, G. and Liu, J.: Human Action Recognition From Various Data Modalities: A Review, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 45, No. 3, pp. 3200–3225 (online), 10.1109/TPAMI.2022.3183112 (2023).
- [4] Nguyen, H.-C., Nguyen, T.-H., Scherer, R. and Le, V.-H.: Deep Learning for Human Activity Recognition on 3D Human Skeleton: Survey and Comparative Study, *Sensors*, Vol. 23, No. 11 (online), 10.3390/s23115121 (2023).
- [5] Apicella, G., GiuseppeD’ Aniello, Fortino, G., Gaeta, M., Gravina, R., Tramuto, L. G.: A Situation-aware Wearable Computing System for Human Activity Recognition, *2022 IEEE Intl Conf on Dependable, Autonomic and Secure Computing, Intl Conf on Pervasive Intelligence and Computing, Intl Conf on Cloud and Big Data Computing, Intl Conf on Cyber Science and Technology Congress (DASC/PiCom/CBDCCom/CyberSciTech)*, pp. 1–7 (online), 10.1109/DASC/PiCom/CBDCCom/Cy55231.2022.9927852 (2022).
- [6] Yang, X., Cao, R., Zhou, M. and Xie, L.: Temporal-Frequency Attention-Based Human Activity Recognition Using Commercial WiFi Devices, *IEEE Access*, Vol. 8, pp. 137758–137769 (online), 10.1109/ACCESS.2020.3012021 (2020).
- [7] Cook, D. J., Crandall, A. S., Thomas, B. L. and Krishnan, N. C.: CASAS: A Smart Home in a Box, *Computer*, Vol. 46, No. 7, pp. 62–69 (online), 10.1109/MC.2012.328 (2013).
- [8] Matsui, T., Onishi, K., Misaki, S., Fujimoto, M., Suwa, H. and Yasumoto, K.: SALON: Simplified Sensing System for Activity of Daily Living in Ordinary Home, *Sensors*, Vol. 20, No. 17 (online), 10.3390/s20174895 (2020).
- [9] Lewis, D. D.: A sequential algorithm for training text classifiers: corrigendum and additional data, *SIGIR Forum*, Vol. 29, No. 2, p. 13–19 (online), 10.1145/219587.219592 (1995).
- [10] Settles, B. and Craven, M.: An Analysis of Active Learning Strategies for Sequence Labeling Tasks, *Proceedings of the 2008 Conference on Empirical Methods in Natural Language Processing* (Lapata, M. and Ng, H. T., eds.), Honolulu, Hawaii, Association for Computational Linguistics, pp. 1070–1079 (online), <https://aclanthology.org/D08-1112/> (2008).
- [11] Caron, M., Touvron, H., Misra, I., Jegou, H., Mairal, J., Bojanowski, P. and Joulin, A.: Emerging Properties in Self-Supervised Vision Transformers, *2021 IEEE/CVF International Conference on Computer Vision (ICCV)*, pp. 9630–9640 (online), 10.1109/ICCV48922.2021.00951 (2021).
- [12] Chen, T., Kornblith, S., Norouzi, M. and Hinton, G.: A simple framework for contrastive learning of visual representations, *Proceedings of the 37th International Conference on Machine Learning, ICML’20, JMLR.org* (2020).
- [13] He, K., Fan, H., Wu, Y., Xie, S. and Girshick, R.: Momentum Contrast for Unsupervised Visual Representation Learning, *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 9726–9735 (online), 10.1109/CVPR42600.2020.00975 (2020).