

加速度データからの機械学習による行動認識

北村 和也* 小高 知宏* 黒岩 丈介* 諏訪 いずみ* 白井 治彦**

A Method to Human Activity Recognition Using Acceleration Data by Machine Learning

Kazuya KITAMURA*, Tomohiro ODAKA*, Jousuke KUROIWA*,
Izumi SUWA* and Haruhiko SHIRAI**

(Received February 3, 2020)

In this paper, we proposed a method to identify human behavior using a 3-axis acceleration sensor of a smartphone. To realize context-aware services such as efficient energy-saving appliance control and elderly monitoring, high-accuracy in-home living activity recognition is essential. We tried to improve recognition accuracy by using deep learning for HAR(Human Activity Recognition).

The proposed methods are CNN(Convolution Neural Network) and lstm(Long short-term memory) methods. An experiment was performed using the HASC dataset to verify the effectiveness of the method. The HASC data set is data of three-axis acceleration. As a result of the experiment, the result using the CNN method was 95.4%, and the result using the LSTM method was 94.3%.

Key Words: Human Activity Recognition, Machine Learning

1. はじめに

スポーツの成績向上や日常生活の支援などを目的として人間の行動の分析が行われている。人間の行動を把握できることで、健康管理アプリケーション、スポーツ支援、高齢者の見守りシステムなど生活を支えるシステムに応用できる。これらのシステムに応用するためには、多種多様な行動の種類を識別し、高精度かつ高速に識別できることが求められる。また行動を理解し、行動内容や意図を把握したうえで適切な支援をする必要がある。

行動や立ち振舞を識別し理解する技術として、人間行動認識技術がある。人間の生活行動を観測し、理解できることで生活の質を向上させることが期待されている。近年、スマートフォンやスマートウォッチ

などの小型センサを搭載した機器が普及したことにより、センサで行動を計測することが容易になり、研究が盛んに行われている。

推定方法に用いるセンサは様々なものがあり、大きく分けてビデオカメラなどを用いて画像処理を使用する方法と加速度センサや角速度センサなどを使用する方法に分けられる。従来の行動認識では、SVM(Support vector machine)、ランダムフォレストなどの機械学習を利用した研究が多い^[1]。

そこで本研究では、スマートフォンの3軸加速度センサのデータから高齢者の見守りシステムや介護者の支援システムに応用できることを目的として、行動認識に深層学習を利用することで認識精度の向上を試みた。

データセットとして HASC(Human Activity Sensing Consortium) コーパスに収録されている停止、ジョギング、歩行、スキップ、階段上がり、階段下がり の6種類の行動データを用いて検証を行った^[2]。

本論文では、2章に行動認識における問題点と対策を示し、3章に提案する認識手法について述べる。ま

* 大学院工学研究科 知能システム工学専攻

** 工学部技術部

* Nuclear Power and Energy Safety Engineering Course,
Graduate School of Engineering

** Technical Division

た、4章では、提案手法の実験を行い、5章で実験の結果を述べる。6章では結果についての考察をし、7章では本研究で提案した手法について総括する。

2. 行動認識における問題点と対策

行動認識は機械による人間行動の自動認識である。近年、スマートフォンやタブレットなどのモバイル端末が多種多様なセンサを内蔵することでセンサで人の行動を計測し、行動に対するデータの取得が容易になっている。センサデータを解析することで医療やスポーツ、さらには日常生活などの幅広い分野で応用可能である。

本研究で扱う行動とは歩く、走る、止まるなどの基本的な行動である。行動情報とはセンサなどから対象とする人間の行動することによって取得できる情報であると定義する。本研究では、その行動情報としてスマートフォンの加速度信号情報に注目して、行動の種類を識別を実現することを目的とする。

2.1 行動認識技術の問題点と解決への取り組み

使用するセンサの種類や特徴量、識別するためのアルゴリズムを決定することで行動認識をすることができる。行動認識を行う場合に考える要素とそれに伴う問題点を以下にまとめる。

- 認識の精度
- マルチユーザー性
- 多様性
- リアルタイム性
- 装着感や拘束性
- 導入や運用のコスト

認識の精度とは行動認識の精度である。先行研究においても精度は90%以上の研究は多く存在するが、重要な操作に使用はしにくい。また、認識エラーとしては、そもそも認識しなかったり、他の動作と間違えて認識したりする誤識別や動作していないのに認識してしまうなどが考えられ、重要なシステムに組み込んでしまうと重大な問題を起こしてしまう可能性がある。

多様性やマルチユーザー性については、認識できる行動数をより多く対応することが好ましい。また、単一の行動だけを識別するのではなく、複数の行動を同時に識別できることが求められる。一般に識別で

きる行動の多さやユーザ数が多くなるほど識別精度は低下する。

リアルタイム性では、一般に認識は動作終了後に過去の一定時間間隔のセンサ情報を使用し特徴量抽出を行う。つまり、実際の動作からコンピュータが認識するまでにはラグが存在し、それが使用するユーザにとってはストレスになる場合がある。このタイムラグをどれだけ減らすことができ、リアルタイム性を確保できることが重要である。

装着感や拘束性は、行動認識では身体に複数のセンサデバイスを付けて計測することが一般的であり、拘束性が問題であった^[3]。しかしスマートフォンの登場でそれらの問題はかなり減少されている。

導入や運用コストは、設置型のセンサや装着型のセンサでは導入するまでの労力やコストがかかっていたが、スマートフォンなど既存のものにセンサが組み込まれたデバイスを用いることによって導入コストは格段に下がった。

2.1.1 機械学習を用いた行動認識手法

ウェアラブルセンサとIoTデバイスの増加に伴い、ユーザの行動データというものをリアルタイムで収集する傾向が高まっている。ロジスティック回帰、決定木、ランダムフォレスト、隠れマルコフモデルなどのアルゴリズムを使用して、行動認識の認識精度で大きな進歩を成し遂げてきた^{[4][5]}。このように、機械学習を用いた行動認識手法は多く存在し例えば、ランダムフォレストを利用した人間行動認識を行っており、複数の独立したランダムフォレストの分類器を統合しより安定し正確で高速な分類器を構築している^[6]。実験結果では、93.44%の精度で19個の行動を正しく認識している。

また、ウェアラブルデバイスでの行動認識に畳み込みニューラルネットワーク(CNN)を多く用いている。CNNは、ローカル情報からグローバル情報への重みを共有することでシーケンス全体をモデル化し、一連の畳み込み演算を通じて階層層で抽象的な機能を抽出し、潜在的な機能をキャプチャするための生データの信号を処理することができる。CNNを用いた研究では行動認識の特徴量を自動的に抽出することで効率的に認識が行えるものを提案している^[7]。畳み込み処理により得られる局所的な特徴やプーリング処理により、積み重ねにより得られるさまざまなスケールの特徴が行動認識において重要であることを強調し、主に画像認識で利用されてきた畳み込みニューラ

ルネットを時系列データから成るセンサデータに対して適用する方法を提案しその有効性を検証した。

3. 提案する行動認識手法

3.1 行動認識における畳み込みニューラルネットワークの利用

一般に、CNN は画像データや動画データ等の多次元データを対象に利用することが多い。これは、画像の局所的な特徴量を学習できるため画像認識や物体認識に有効であることが知られている。先行研究では、行動認識においても親和性が高く、認識精度も高いことが示された^[7]。しかし、ニューラルネットへ信号を入力するたびに 2 次データに変換する計算コストが必要となり、計算能力の低い端末の場合、実際の運用には適さない可能性がある。

そこで、1 次信号を 2 次信号に変換することなく入力する 1D CNN に着目した。1D CNN の畳み込み層は入力と同じく 1 次元の畳み込みフィルタを用いて時間方向へスライドさせる演算を行う。基本的には通常の CNN と同様に実装することができる。

3.2 行動認識におけるリカレントニューラルネットワークの利用

上記の CNN は、局所的な周波数帯域での特徴量に適合している。本研究で扱う行動データは 3 軸の加速度信号データであり時系列データである。時系列データに対する深層学習では時間関係の特徴を抽出することができる RNN(リカレントニューラルネットワーク)よく用いられる。また、LSTM(Long Short Term Memory) は、過去のタイムステップを長期間保持することができる特性があり、この LSTM を用いることでより精度のよい行動認識モデルが期待できる。

4. 実験

本節では、提案した手法の実験を行う。実験にともない用いるデータセットと深層学習モデルについて述べる。

4.1 用いるデータセット

この節では、使用するデータセットの説明をする。データセットは、スマートフォンによって収集された 3 軸加速度センサを用いたデータである HASC コーパスを用いる。

このデータセットの対象としている行動は、停止(stay)、歩行(walk)、ジョギング(jog)、スキップ(skip)、

階段を上がる(stUp)、階段を下る(stDown)の 6 種類であり、それぞれの 3 軸加速度信号データ(csv)、メタデータ(meta)、ラベルデータ(label)の 3 種類のデータ形式が記録されている。また、データセットは、大きく分けてセグメントデータとシーケンスデータに分けることができる。

セグメントデータは対象とする 6 種類の行動に対して一つ一つ記録しているデータである。各 1 回の計測時間は 20 秒でありサンプリング周波数は 100 Hz のものを用いる。1 人あたりの計測回数を 5 セットとしている。データそれぞれ加速度データにメタデータが付与されており、メタデータには使用デバイス名、サンプリング周波数、性別、行動、身長、体重、スニーカー・サンダル・革靴などの履物、アスファルト・フローリングなどの路面状態、廊下などの場所、腰ポケット・ベルト・首・胸ポケットなどのセンサ取り付け位置、センサが固定なのか可動なのかのセンサデバイス取り付け状態の記載がされている。また、ラベルデータも付与されている。

シーケンスデータは、6 種類の行動すべてが含まれた行動データである。各行動はそれぞれ 5 秒以上続けられておりそれが連続して行動されてひとつのデータとなっている。1 回の計測時間は 120 秒であり、セグメントデータ同様に加速度データにメタデータ、ラベルデータが付与している。

本研究では、本データセットのセグメントデータを用いて実験を行う。セグメントデータの詳細を表 1 に示す。

4.2 実験に用いるモデルの構築

深層学習を用いて行動認識を行う。用いる深層学習モデルは、1 次元の畳み込みニューラルネットワークである 1D-CNN と RNN の拡張として登場した、LSTM を用いて深層学習モデルを構築した。

4.2.1 畳み込みニューラルネットワークによる行動認識の流れ

本研究では 1 次元の CNN である 1D-CNN を用いて行動認識を行う。

1 次元 CNN は、2 層の畳み込み層と 1 層の全結合層から構成されている。全ての畳み込み層において、1 次元重みフィルタを横軸方向のみに移動して畳み込み処理を行う。それにより、各行動データごとの特徴からそれぞれの行動を識別する。1 次元の畳み込みウィンドウの長さは 3 で元の入力と同じ長さを出力がもつように入力にパディングを行っている。

表 1 HASC セグメントデータの詳細

セグメントデータ	
一回の計測での信号	時刻 (sec)、 X 軸 (G)、 Y 軸 (G)、 軸 (G)
サンプリング周波数	20 ~ 100 Hz
対象とする行動	“ stay ”, “ walk ”, “ jog ” “ skip ” “ stair up ”, “ stair down ”
1 回の計測時間	20 s
被験者数	7 人
試行回数	540
Type	Single activity

表 2 RNN の設定

設定項目	設定値
出力層の活性化関数	ソフトマックス関数
誤差関数	クロスエントロピー関数
中間層ユニット	LSTM
最適化手法	Adam
入力次元数	3
出力次元数	6

入力データに対して 1 次元重みフィルタを畳み込みを行い特徴量を得る。次に、1 次元畳み込み層では、前の層目で取得した特徴量に対して 1 次元重みフィルタを畳み込み、新たな特徴量を得る。そして、MaxPooling を行う。最後に全結合層に与え出力層で softmax 関数を用いて算出する。

4.2.2 リカレントニューラルネットワークによる行動認識の流れ

3 軸加速度信号を入力として、各時刻の 3 軸加速度データを 3 次元の入力層と行動のクラス数に対応する出力層を構築した。中間層の各ユニットは LSTM を結合したネットワークモデルを用いる。出力層の活性化関数はソフトマックス関数を使用し、活性化関数にはクロスエントロピー関数を使用する。最適化手法は Adam を用いた。この RNN の設定の詳細を表 2 に示す。このネットワークは入力ベクトルが入力されたときに得られた出力ベクトルの要素の中で最大値を持つ要素に対応する行動を出力ベクトルとする。

5. 結果

本章では、4.1 節 で述べた HASC データセットを用いて実験を行い、CNN と LSTM を用いた手法の結

果を述べる。

5.1 畳み込みニューラルネットワークを用いた手法の結果

1 次畳み込みニューラルネットワークを用いて行動認識を行った結果を以下に示す。混合行列を表 5.1 に示す。全体の認識精度としては、95.4% となった。

最大の認識率は Stay であり、他の行動と間違えることがない。その次に認識精度が良かったのが、skip でありこれも他の行動状態と比べて行動が特徴的であるため他の行動より認識精度が高い。反対に精度が悪かった行動は、階段を下るである Stair down であり、Walk や Stair down などの他の行動と誤認識していることがわかる。

5.2 リカレントニューラルネットワークを用いた手法の結果

RNN を用いて行動認識を行った結果を以下に示す。RNN のモデルとして以下のパラメータとなった。設定したパラメータは、中間層である LSTM の数、遡るタイムステップ数、ドロップアウト率である。中間層の数は 3、遡るタイムステップ数は 30、ドロップアウト率は 50% に設定をした。表 4 に設定したパラメータをまとめる。

LSTM で行った分類結果の混合行列を表 5 に示す。全体での認識精度としては、94.3% となり、上記の CNN で行った識別とほぼ同値である。最大となった行動状態は Stay であり、その次に Skip の 94.2% である。最低の認識結果となった行動状態は、Stair up であり、CNN の場合は Stair down であったため違う結果となった。

表 3 CNN による分類結果

	Jog	Stair down	Skip	Stair up	Stay	Walk
Jog	0.950	0	0.004	0	0	0
Stair down	0.017	0.906	0.008	0.018	0	0.020
Skip	0.016	0.002	0.977	0.004	0	0
Stair up	0.017	0.045	0.006	0.944	0	0.028
Stay	0	0	0	0	1	0
Walk	0.001	0.044	0.006	0.023	0	0.950

表 4 RNN を用いた最良のモデル

パラメータ	値
中間層の数	3
遡るステップ数	30
ドロップアウトレート	0.5

6. 考察

CNN と LSTM の 2 つの手法を用いて行動認識を行ったがこの 2 つの手法は、ほとんど同様の認識結果になったことが結果からわかる。少しの差であるが、CNN を用いた手法のほうが 1% 程度高くなっている。どちらの手法も全体として 90% 以上の結果であるため有効であると言える。どちらの手法も stay の状態に対してはほとんど間違えることはない。

CNN を用いた手法は、全体として 95.4% となった。それぞれの行動について混合行列で見ると、Stair down が Walk がそれぞれ誤認識している。Jog は他の行動それぞれと少しずつ誤認識していることがわかる。

RNN を用いた手法は、全体として 94.3% になった。混合行列のそれぞれの行動についてみると、Walk が認識結果が良くない結果になっている。Walk は他の行動とそれぞれ誤認識しており、中でも Stair up や Skip に誤認識していることがわかる。また、Jog と Stair down と Skip と Stair up の間でも誤認識しあっている。歩くや走る行動状態の時、重力方向に動きが大きいとご認識しやすいと考える。

誤認識しているそれぞれの行動の元データを見ると、一般にあまり重力方向にブレない行動のとき大きなブレのある行動をしていると他の行動と誤認識していることがわかった。ここに注目して、モデルを作ればもう少し精度が高くなったと考えられる。

誤認識を少なくするには、3 軸の加速度データだけでなく、他の角速度データや音データなどを併用す

ることで誤認識は減り、認識精度が高くなったと考える。

健康管理や見守りシステムに利用することをすると、ユーザが正常な生活ができていないかを判断するかが重要である。本研究で使ったデータセットである HASC コーパスは基本的な動作しか対象にしている。見守りシステムや健康管理アプリケーションに必要なのは日常生活の監視であるため他のデータセットやデータセットの拡張が必要である。

7. まとめ

本章では、加速度データからの機械学習による行動認識に提案した手法について総括し、今後の課題点について述べる。

本研究では、健康管理アプリケーションや高齢者の見守りシステムなどの生活を支えるシステムに応用される技術である行動認識技術についてスマートフォンから取得できる 3 軸加速度データを利用して分類を行いました。介護や高齢者の見守りシステムに応用できることを目的として加速度信号データから日常行動の識別を目的とした手法の提案および実装を行いました。

本研究で提案した手法は、深層学習である 1 次元の CNN である 1D-CNN を用いた手法と、LSTM を用いた手法の 2 つの手法を提案した。提案手法の有効性を検証するため人間行動理解のための装着型センサによる大規模データ構築を行っている HASC コーパスのスマートフォンから取得した 3 軸加速度データを用いて実験を行った。実験の結果、CNN を用いた手法では、95.4% であり、LSTM を用いた手法では 94.3% という結果となった。

今後の課題として以下の点が挙げられる。

はじめに、提案した手法の改善が必要である。本研究での最良の認識結果であるが 95.4% であるが、これは約 5% も誤認識をしているということである。介

表 5 LSTM による分類結果

	Jog	Stair down	Skip	Stair up	Stay	Walk
Jog	0.939	0	0.021	0	0	0
Stair down	0.017	0.936	0.011	0.027	0	0.023
Skip	0.016	0.012	0.942	0.021	0	0.019
Stair up	0.017	0.032	0.010	0.926	0	0.023
Stay	0	0.003	0	0.004	0.993	0.009
Walk	0.011	0.018	0.015	0.022	0.007	0.927

護者の支援や見守りシステムなどの基幹システムとして運用した時に、深刻な被害に繋がる可能性もある。そのため、最終的な目標としては、100%の分類精度を目指す必要がある。

本研究では、3軸加速度データを用いて行動分類を行い、行動認識に3軸加速度データを用いるのは有効だといえる。精度をよりよくすることとしては加速度データ以外のデータと併用して識別することでより良くなると考える。本研究はスマートフォンから取得した加速度データを用いたので、スマートフォンに搭載されているセンサから取得するのが妥当であろう。角速度データや音データを学習時に併用することができれば、更に用意するセンサデバイスも必要なくより良い精度が期待できる。

加えて、データセットについてである。今回用いたデータセットは6行動のデータであった。実環境における人間の行動分類としては非常に少ない。本研究の手法は、6行動の分類での分類結果でありさらに識別する行動数が増えれば精度は低下するだろう。介護や見守りシステムや健康管理アプリケーションへの応用を考えた時、重要であるのは健康に関わる行動をしているのかどうかである。それぞれの行動に対して高度の重要度を設定しアプリケーションに応用するなどの改善が必要である。

次に、分類を行ったアルゴリズムについてである。本実験ではCNNとRNNを用いたが、これら以外にも様々なアルゴリズムが存在する。他の機械学習アルゴリズムを用いて劇的な改善とはならないが多少の改善が期待される。また、今回用いたアルゴリズムを繋げてひとつの機械学習モデルにすることでよりよい結果になることも期待できる。

上記の課題となる事項を取り入れることにより、高い精度で行動認識が可能になると期待できる。

参考文献

- [1] Ling Bao and Stephen S Intille. Activity recognition from user-annotated acceleration data. In *International conference on pervasive computing*, pp. 1–17. Springer, 2004.
- [2] Nobuo Kawaguchi, Nobuhiro Ogawa, Yohei Iwasaki, Katsuhiko Kaji, Tsutomu Terada, Kazuya Murao, Sozo Inoue, Yoshihiro Kawahara, Yasuyuki Sumi, and Nobuhiko Nishio. Hasc challenge: gathering large scale human activity corpus for the real-world activity understandings. In *Proceedings of the 2nd augmented human international conference*, pp. 1–5, 2011.
- [3] 佐藤誠, 森田千絵, 土井美和子ほか. 生体データと加速度データを用いた行動認識. 第65回全国大会講演論文集, Vol. 2003, No. 1, pp. 105–106, 2003.
- [4] Jennifer R Kwapisz, Gary M Weiss, and Samuel A Moore. Activity recognition using cell phone accelerometers. *ACM SigKDD Explorations Newsletter*, Vol. 12, No. 2, pp. 74–82, 2011.
- [5] Francisco Javier Ordonez, Gwenn Englebienne, Paula De Toledo, Tim Van Kasteren, Araceli Sanchez, and Ben Kröse. In-home activity recognition: Bayesian inference for hidden markov models. *IEEE Pervasive Computing*, Vol. 13, No. 3, pp. 67–75, 2014.
- [6] Zengtao Feng, Lingfei Mo, and Meng Li. A random forest-based ensemble method for activity recognition. In *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pp. 5074–5077. IEEE, 2015.
- [7] Jianbo Yang, Minh Nhut Nguyen, Phyo Phyo San, Xiao Li Li, and Shonali Krishnaswamy.

Deep convolutional neural networks on multichannel time series for human activity recognition. In Twenty-Fourth International Joint Conference on Artificial Intelligence, 2015.

