

## 深層学習による日常健康データに基づく高齢者メンタル変調検知 Mental depression detection of the elderly based on daily health data using deep learning

杉本 千佳<sup>†</sup>      大久保 遼河<sup>‡</sup>  
Chika Sugimoto   Ryoga Okubo

### 1. はじめに

高齢者は環境の大きな変化に伴う影響を受けやすく、コロナ禍での自粛生活長期化により、心身機能低下や精神面の不調が進行する恐れが危惧されている。このため、心身の変調の予兆を早期に捉えケアを行うことが、健康寿命を延伸し QOL を改善するために必要となっている。多くの高齢者が介護サービスや医療ケアの提供を受ける介護施設では、持病や心身の状態が個々に異なる高齢者の心身の変調を早期に検知してケアにいかすことが重要である。一方で、介護施設は人手不足が課題となっており、介護職員の負担軽減や介護サービスの向上のために ICT 活用が期待される<sup>[1]</sup>。ICT 技術を活用し、センサでバイタルサインや行動を把握しリスク予知を行うことで、高齢者の健康状態や安全を遠隔で人的コストを抑えて効率的に確認することが可能になり、そのためには正しくリスクを予測する手法が必要である。そこで本研究では、介護施設に通う高齢者から取得された日常生活における活動や睡眠、バイタルなどの健康データをもとに、深層学習を用いてメンタル状態を推定するモデルを構築し評価した。

### 2. 手法

ストレスに関連するパラメータとして、安静時心拍数、心拍変動指標、皮膚温度、呼吸数、睡眠レベル・時間、活動量（歩数）等のデータとそれらの統計量を特徴量候補とし、長期ストレススコアと相関が高い特徴量を選択する<sup>[2]</sup>。各特徴量は 1 日単位の値として算出し、個人毎に正規化して入力データとする。推定モデルは、時系列データの学習に適した深層学習の LSTM を用いて構築した。LSTM モデルは、単層と多層、順方向と逆方向、シングルパスとマルチパスを組み合わせて設計して精度の高低を評価し、最適なアーキテクチャを求めた（図 1）。事前検討として、個別のデータからストレス推定に重要度の高い特徴量を評価したところ個人毎に差異が見られた。そこで、推定モデルは共通アーキテクチャモデルをもとに個人のデータのみで個別に学習させて個人毎に構築することとした。連続する 7 日間のデータを入力データとし、その翌日のデータを入力データとして推定する。

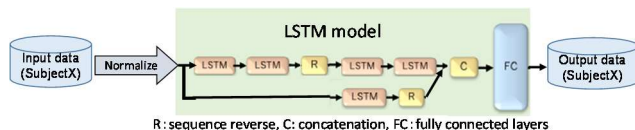


図 1 推定モデル例

<sup>†</sup> 横浜国立大学大学院工学研究院 Faculty of Engineering, Yokohama National University

<sup>‡</sup> 横浜国立大学大学院理工学部 Graduate School of Engineering Science, Yokohama National University

### 3. 実験

コロナ禍においてデイサービスまたはデイケアを提供する介護施設に通う高齢者の長期ストレス状態を推定するため、高齢者が日常生活の中で装着可能なウェアラブルデバイスを用いて毎日の健康データを取得し、収集したデータをもとにモデルを構築した。なお、本研究および実験は、横浜国立大学の倫理審査の承認を得て、社会実験実施母体の倫理審査の承認を経て実施した<sup>[1]</sup>。

#### 3.1 データ

社会実験に参加した介護施設に通う被介護者 17 名（男性 9 名、女性 8 名）の約 2 か月間のデータを解析対象とした。事前に実験の目的及び手順を説明し、参加への同意を得たうえでデータの収集を行った。実験期間中、各自の日常生活においてスマートウォッチ Fitbit sense<sup>[3]</sup> を装着してもらうことでデータ取得される。本実験において Fitbit Sense で収集されたデータは、表 1 に示すバイタルデータ、歩行やエクササイズなどのデータ、睡眠の質に関するデータなどの 18 種類である。また、こころのストレス度チェック<sup>[4]</sup> をもとに作成したメンタル問診票に 1 週間に一度定期的に回答してもらうことで、メンタルの主観評価値を取得した。メンタル問診票は 10 項目の質問から成り、6 項目の 5 段階評価の回答をもとに 0~24 点でメンタルスコアの合計点を算出し、続く 2 項目の二者択一の回答をもとにストレスレベルを判定する。残りの 2 項目はストレスの原因や対処法に関する質問のため、それらの回答結果は本評価では利用しない。Fitbit sense のストレススコアデータ（0~100 でスコア化）は入力データとしては用いず、メンタル問診票から得られたメンタルスコアのデータを補完するためと、正解データの一つとして使用するために用いる。メンタル主観評価値を取得できるのは 1 週間に一回だけのため、残りの 6 日間分をストレススコアの変動値をもとに補完してメンタルスコアの正解データとした。これにより、正解データを主観評価結果とした場合とバイタル等データからの評価結果とした場合の 2 通りを想定して検証する。

また本実験では、収集した 17 名のデータの内、1 週間に 1 回メンタル問診票の回答を得られ、かつ Fitbit sense のデータ欠損数が一定以下の人のデータをモデル構築の対象とした。得られたデータを個別におおよそ 8:2 の割合で訓練データとテストデータに分け、訓練データで学習を行った後テストデータで評価した。

表 1 主な取得データ

| 項番 | 取得データ   |
|----|---------|
| 1  | 安静時の心拍数 |
| 2  | アクティブ時間 |
| 3  | 座っている時間 |
| 4  | 心拍数     |

|    |             |
|----|-------------|
| 5  | 歩数          |
| 6  | カロリー        |
| 7  | 高度          |
| 8  | 距離          |
| 9  | 手首温度        |
| 10 | SpO2(睡眠中)   |
| 11 | 睡眠レベル       |
| 12 | 睡眠スコア       |
| 13 | 心拍変動(睡眠中)   |
| 14 | 呼吸数(睡眠中)    |
| 15 | 推定酸素変動(睡眠中) |
| 16 | 睡眠の記録       |
| 17 | エクササイズの記録   |
| 18 | ストレススコア     |

### 3.2 モデル

モデルアーキテクチャは、層数を 1~6 層として順 LSTM 層と逆 LSTM 層の組合せをかえて 3 パスまでの組合せについて評価した。また、過学習を防ぐために各層に Layer Normalization を適用した。順 LSTM 層、逆 LSTM 層を連結して多層化、マルチパス化したアーキテクチャのモデルを構築し個人毎に精度比較した結果、最高精度を達成する個人毎の最適モデルは同一なアーキテクチャモデルではなかった。しかし、共通して高い精度を達成したアーキテクチャがあったため、そのモデルを共通モデルとして採用した。アーキテクチャがシンプルな構成の方が学習において収束しやすく、汎用的に用いるには有用であった。5 層以上にすると、過学習によりエポック数を増やすと収束しにくくなり誤差が増加する傾向が見られた。よって、今回共通モデルとして採用した LSTM モデルは、複数の順方向と逆方向の LSTM 層を連結させたシングルパスのモデルである。

### 3.3 評価指標

推定精度の評価指標として、以下の式に示す平均絶対誤差(MAE)を用いた。

$$MAE = (1/n) \sum_{i=1}^n |f_i - y_i|$$

( $n$ :データ数,  $f_i$ :推定値,  $y_i$ :実測値)

各モデルについて 10 回ずつ評価し、テストデータによる MAE の平均値をそのモデルの精度とする。

## 4. 結果と考察

本実験では 5 名分のデータに対してモデル構築を行い精度評価した。長期ストレススコアを推定した精度結果を表 2 に示す。また、参考値として最良値も示す。図 2 に示す MAE の例のように、学習は安定し早期に収束した。各層に Layer Normalization を適用することで、適用しない場合と比較して収束が早くなり、かつ精度が一番良くなる結果となった。一方で、いずれのモデルでもテスト時の MAE が訓練時と比較しかなり大きくなった。理由として、学習

表 2 平均絶対誤差の平均値と最良値

| 正解データ   | Average_test_MAE | Best_test_MAE |
|---------|------------------|---------------|
| メンタル問診票 | 0.227            | 0.134         |
| ストレススコア | 0.186            | 0.078         |

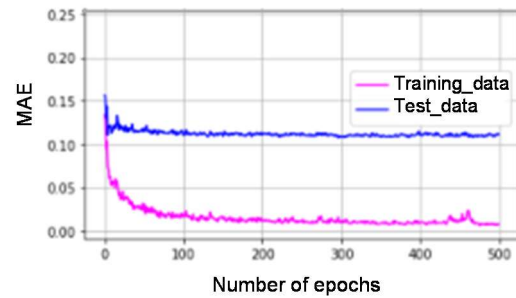


図 2 平均絶対誤差の結果例

データ数が十分ではなかったために、モデルの汎化性能が十分得られなかったことによると考えられる。

また、正解データとしてストレススコアを用いた場合の方が主観評価値を用いた場合より精度が高くなった。これは高齢者が必ずしも自身のメンタル状態を正確に認識できていない可能性が考えられる。

## 5. おわりに

LSTM を用いて高齢者に対して汎用を使用できる共通モデルを導出し、日常的に取得が可能な時系列バイタル等データをもとに長期ストレス状態を推定する手法を提案した。重要な特徴量や最適モデルに人による差異は見られたものの、共通モデルを用いても従来手法と比較し誤差を低減してストレスレベルを推定できることが示唆された。本実験では 17 名のうち 5 名分しか評価に適したデータを得られなかったため、対象を高齢者とした場合のデータ取得方法と解析方法の更なる工夫が必要である。

今後は更なる推定精度の向上を図るためにデータの質を高めることが必要であり、解析対象の人数とデータ数(期間)を増やしてデータサンプル数を多くしたうえで最適モデルを導出し、汎用モデルとした場合の有効性検証を行う。これにより、高齢者を対象とした場合の長期ストレス状態の把握とストレス要因の分析に寄与する。

### 謝辞

本研究成果は、国立研究開発法人情報通信研究機構の委託研究により得られたものである。また、本研究にご協力頂いた介護施設の皆様および社会実験を実施いただいた富士通様に感謝の意を表す。

### 参考文献

- [1] 渡邊祥太, 葛谷雄一, 長島久美子, 長坂知美, 杉本千佳, 小林匠, 河野隆二, “機械学習を活用した体調変動検知における介護現場感染症対策の取り組み”, 信学技報, vol. 121, no. 404, MICT2021-106, pp. 29-33 (2022).
- [2] 岡崎裕子, 杉本千佳, “COVID-19 下でのウェアラブルセンサーデータに基づく長期ストレス推定”, 電子情報通信学会総会大会, D-20-13, (2022).
- [3] Fitbit Sense, <https://www.fitbit.com/global/jp/products/smartwatches/sense>, (参照 2022-6-10).
- [4] 内閣官房, こころのストレス度チェック, <https://corona.go.jp/selfcheck/>