

# 日常生活の中の生体情報計測と解析による

## 心身状況の把握に関する基礎的検討

生体計測工学研究室 2181212 江原 和

(主査：京相 雅樹 教授 副査：和多田 雅哉 教授, 小林 千尋 講師)

### 1. はじめに

近年、精神疾患により医療機関にかかっている患者数は大幅に増加しており、平成 26 年では 392 万人、平成 29 年では 419 万人となっている<sup>[1]</sup>。特に、うつ病の患者数は著しく増加している。平成 30 年の労働安全衛生調査によると現在の仕事や職業生活に関することで、強いストレスになっていると感じられることがある労働者の割合は 58%となっている<sup>[2]</sup>。過度なストレスはうつ病などの精神疾患のリスクを高めるだけでなく、免疫力を低下させ感染症のリスクを上げる、循環器系や消化器系に障害を与えるなど様々な疾患の原因となる。

ストレスは大きく 2 つの種類に分類される。1 つはライフイベントなどの心身に大きな影響を与える出来事によって発生するストレス、もう一つはデイリーハッスルと呼ばれる日常生活の中の出来事によって発生する短期的なストレスである<sup>[3]</sup>。

ライフイベントによって発生するストレスはその出来事自体を回避することが難しく、個人の心がけなどでは解決が難しくソーシャルサポートなどが必要になる。一方で、デイリーハッスルは些細な出来事が原因であり、その原因は短時間で消えてしまうことが多い。デイリーハッスルの発生原因については個人の特性的な問題、環境的な問題などいくつかの原因が挙げられるが、どれも自覚することが難しい。よってデイリーハッスルによるストレスはそれに起因する症状が出るまで本人に知覚され辛く、症状の悪化、または無自覚に日常生活に悪影響を及ぼしている可能性も考えられる。このようなことから、日常的な生体計測からストレスを計測し、そこから健康管理を行うことは精神疾患の予防、生活の質の向上に役立つと考えられる。また、日常生活での一過性のストレスを連続的に計測することで自力では難しいデイリーハッスルの原因の特定、また、その軽減が可能になると考えられる。

現在、心拍変動を用いたストレス評価法が研究されており、ウェアラブルデバイスを用いたストレス計測についても研究が行われている<sup>[4]</sup>。しかし、連続的なストレス計測を可能とするウェアラブルデバイスで実用化されているものは少ない。そのため日常生活におけるストレスを検知し、ストレスサーの特定などに役立てることは難しい。連続的なストレス計測を可能にするためには、計測機器がウェアラブルであり、常に計測をおこなえる必要があるほかに短時間間隔でストレス評価を行う必要がある。

本研究ではデイリーハッスルの特定やそれが発生しやすい状況などの特定を目的とした日常生活中での一過性のストレス評価手法について検討する。修士の研究では腕時計型ウェアラブルデバイスから瞬時心拍数を取得し解析を行うまでのシステムを構築した。瞬時心拍数からは自律神経指標を算出しそれを特徴量とし、同時に質問紙法によるストレスの主観評価を行いその結果をもとにラベル付けし、機械学習を行った。

## 2. 理論

### 2.1 HRV 指標

心臓の拍動間隔（RRI）は一定ではなく変動しており、その変動は HRV と呼ばれている。この心拍ゆらぎにあらわれる周期的な変化には、様々な生体のフィードバックメカニズムが関係している。そのなかで、自律神経に大きく関わるものは、呼吸と同じ周期をもつゆらぎと、血圧の変動と同じ周期をもつゆらぎの2種類である。自律神経が正常に機能している場合、呼吸と同期したゆらぎと、血圧と同期したゆらぎの両方が心拍間隔のゆらぎとして観察される。交感神経活動が活発になると HRV は低くなり、副交感神経活動が活発になると HRV が高くなることが知られている。このようなことから自律神経の状態をリアルタイムに、かつ、人体に負担の少ない状態で計測できる可能性をもっており、様々な応用が検討されている。SDNN は RRI（NN 間隔とも呼ばれる）の標準偏差であり、RRI の全体的な変化を表す指標とされている。RMSSD は連続して隣接する RRI の差の2乗の平均値の平方根であり、迷走神経緊張度の指標とされている。またこれらの比である SDNN/RMSSD は交感神経系の活動を表すことができる。

### 2.2 POMS 短縮版<sup>[5]</sup>

ネガティブな気分状態を評価する質問紙形式の方法である。POMS と比較すると質問項目は半分程度となっている。緊張、抑うつ、怒り、活力、疲労、混乱の6つの尺度とそれぞれの尺度につき5項目の質問項目があり、「まったくなかった」(0点)~「非常に多くあった」(4点)までの5段階(0~4)で回答してもらうことで回答者の気分状態を把握することができる。POMS 短縮版は各尺度で5項目の合計得点を算出し T 得点を算出することで評価が行われる。T 得点は以下の式で求められる。

$$T \text{ 得点} = 50 + 10 \times (\text{素点} - \text{平均値}) / \text{標準偏差}$$

これは平均が 50、標準偏差が 10 の正規分布に近似するように変換した値であり、平均点、標準偏差については年代、性別ごとに統計が取られているものを回答者に応じて当てはめる。

### 2.3 ニューラルネットワーク (多層パーセプトロン)

ニューラルネットワーク(NN)は人工知能分野におけるアルゴリズムの一つであり、人間の脳の構造を模して造られている。NN はノードと入力層、隠れ層、出力層から構成されており、前段の層と後段の層にはノード同士のつながりを表す重みでつながっている。この重みは、NN による学習結果と与えられた教師データ（正解ラベル）を用いて重みを徐々に更新して正しい分類が行えるように学習を進める。

## 3. 実験方法

### 3.1 計測方法

日常生活中、被験者 3 名(20 代、男性)にウェアラブルデバイス(fitbit versa3)を装着してもらい、連続的に生体情報を計測した。ウェアラブルデバイスの仕様を Table1 に示す。計測された生体情報は fig.1 のようにデバイスから bluetooth 通信によりスマート

フォンへ、そこから websocket 通信により server へ送信される。また、生体情報の計測中に POMS 短縮版によるストレス主観評価を行ってもらった。ストレス主観評価についてはストレスを受けた直後及び、ランダムなタイミングで回答してもらった。ストレスを受けた瞬間の生体情報を取得するため、ストレスの主観評価はグーグルフォームを使用し、スマートフォンから手早く回答できるようにした。主観評価は回答する作業自体がストレスとして認識されること、回答中の時間経過によりストレス主観評価と精神状態がずれてしまうことを考慮し質問項目が少なく短時間で気分状態が評価できる POMS 短縮版を使用した。

Table1 Specifications of fitbit versa3

Sensor	Infrared sensor 3-axis accelerometer Altimeter Ambient light sensor Optical heart rate sensor Device temperature sensor
Battery life	More than 6 days
Size(W × D × H)[mm]	40.48 × 40.48 × 12.35
Heart rate data storage interval [s]	1
Communication method	Bluetooth



Fig.1 Data communication method

### 3.2 解析方法

本研究では NN によるストレスの有無の判別、ストレスレベルの推定の 2 パターンの解析を行った。ストレスの有無の判別は 2 値分類であり、ストレスレベルの推定は回帰分析である。

#### 3.2.1 正解ラベル

正解ラベルには POMS 短縮版により得たストレス主観評価結果を用いた。POMS 短縮版は 6 尺度 30 項目の回答項目があるが本報告では精神的ストレスと強く関係する緊張、怒り、抑うつ の 3 尺度 15 項目をもとに正解ラベルを設定した。ストレスの有無の判別では緊張、怒り、抑うつ の 3 尺度のうち一つでも T 得点が 50 以上になった場合の正解ラベルをストレスあり(1)、それ以下の場合の正解ラベルをストレスなし(0)としてラベル付けした。ストレスレベルの推定では 60 点以上を(4)、55 点以上 60 点未満を(3)、50 点以上 55 点未満を(2)、45 点以上 50 点未満を(1)、45 点未満を(0)としてラベル付けした。

#### 3.2.2 特徴量

主観評価時刻の生体情報を取得することで、同じ時間帯の主観評価と瞬時心拍数が

一組になるようにした。瞬時心拍数は主観評価時刻の1分前から回答時刻までの1分間を1Hzで取得した。特徴量として最適な解析データの時間幅を特定するため、同様に2、3、4、5分間のデータも取得し解析を行った。それぞれの長さの瞬時心拍数から平均値、分散値、SDNN、RMSSD、SDNN/RMSSDを算出しこの5つを特徴量とした。

### 3.2.3 データセット

被験者ごとに正解ラベル、特徴量の組の70%を訓練データ、30%をテストデータとして分割した。ストレスの有無の判別におけるデータの正解ラベルは、被験者によってはストレスなし(0)が大半を占める偏りのあるデータとなっている。この場合、正解ラベル(0)の状態を過度に学習してしまい、適切な判別が行えないことが知られている。そのため、訓練データにのみデータ拡張を行い、正解ラベルに偏りが無い訓練データを用意した。データ拡張は不均衡データの少数派データを増やす **oversampling** の一種である SMOTE(Synthetic Minority Oversampling Technique)を用いた。訓練データおよびテストデータは正規化処理を行い最終的なデータセットとした。

### 3.2.4 識別処理と評価手法

識別には5層NNを用いた。ストレスの有無の判別、ストレスレベルの推定に使用したNNの構成をTable 2に示す。ストレスの有無の判別では、正解率(accuracy)、精度(precision)及び再現率(recall)を算出してモデルの性能を評価した。それぞれの指標は10回学習とテストを行った際の平均値となっている。精度はストレスあり(1)と分類した内、実際に正解であった割合を表す。再現率は実際にストレスあり(1)とラベル付けされたデータがどれだけ正解に分類されたかを表す。これらの値からストレス識別のモデルを評価することができる。二値分類における混同行列の算出方法をTable 3に示す。ストレスレベルの推定では、推定結果と正解ラベルの散布図、相関係数を求め評価指標とした。

Table 2 NN condition

Parameter	Distinction	presumption
Input layer size	3,5	3
Hidden layer size	64,32,64	64,32,16
Activation function	ReLU	ReLU
Output layer size	2	1
Loss function	Sparse categorical crossentropy	Huber loss
Epoch	100	600
Optimization method	Adam	adam

Table 3 Confusion matrix

		predicted	
		positive	negative
true	positive	TP	FN
	negative	FP	TN

True Positive (TP) ,False Positive (FP) ,True Negative (TN) ,False Negative (FN)  
 $accuracy = (TP+TN/TP+FP+TN+FN)$  , $precision = (TP/TP+FP)$  , $recall = (TP/TP+FN)$

#### 4. 結果・考察

##### 4.1 異なる時間幅でのストレス状態の判別

時間幅の異なる5パターン(1,2,3,4,5分間隔)のデータセットを用いてストレス状態の識別を行い算出した正解率の結果を Fig.2 に示す。識別には5つの特徴量すべてを使用した。被験者1では0.6~0.7程度の正解率、被験者2では0.7~0.9程度の正解率、被験者3では0.5~0.6程度の正解率が得られた。時間幅ごとの正解率を比較すると被験者1、被験者3では3分間のデータ、被験者2では2分間のデータが最も高い正解率となっている。時間幅がそれより短くなる場合、長くなる場合ともに正解率は低下する傾向にある。解析に使用するデータの時間幅が短すぎる場合、ストレス状態の識別に必要な情報が不足するため正解率が低下することが考えられる。また、時間幅が長すぎる場合、主観評価と異なる心理状態の瞬時心拍数まで取得してしまうため、主観評価と瞬時心拍数が一致せず正解率が低下していると考えられる。被験者ごとに正解率には差が見られるが、上記の傾向はどの被験者にも見られた。この結果から主観評価を行った時刻から2~3分ほど前までの瞬時心拍数がストレスの影響を最も反映していると考えられる。

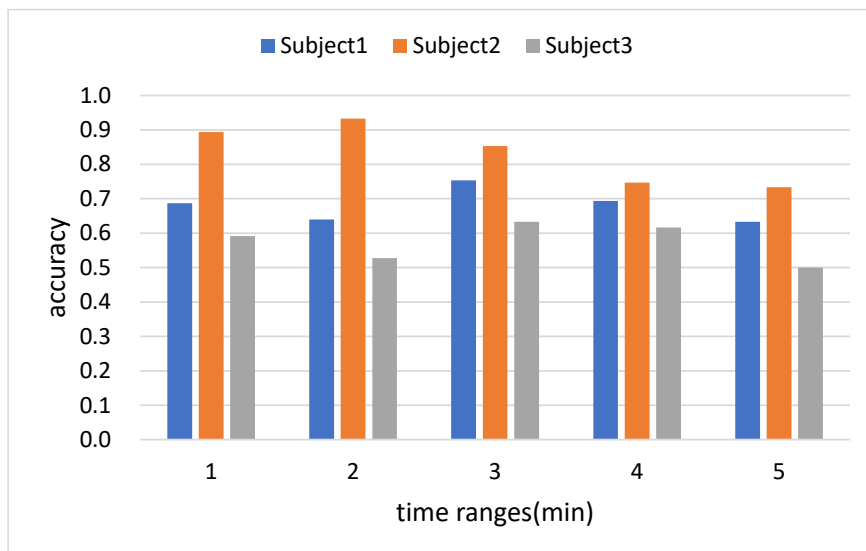


Fig.2 異なる時間幅でのストレス判別の正解率

算出した精度および再現率のうち被験者 1 のものを Table 4 に示す。算出された結果はおよそ正解率と同様の結果となっている。再現率は特に低い値となっており、ストレス状態を見逃す可能性が高いことが分かった。これは学習データにおけるラベル付けの閾値に問題があると考えられる。本研究では T 得点が 50 以上のものをストレスあり、それ未満のものをストレスなしとラベル付けした。しかし、閾値付近のデータが全データの半数近くを占めるため、判別が難しくなっていると考えられる。それに加え、ストレスなしのデータが全体の 3 分の 2 ほどとストレスありのデータより多くなっていることから、閾値付近のデータについてストレスなしの判別結果が増え結果として再現率の低下につながっていると考えられる。今後の検討としては学習データにおいて、ストレスありとなしのデータの T 得点の差を大きく設定することが考えられる。また、そのためにデータ数を増やすこと、特に再現率の向上を目指し強いストレスを受けた際のデータを用意する必要がある。

Table 4 被験者 1 における精度と再現率

	precision	recall
1min	0.742	0.643
2min	0.555	0.464
3min	0.733	0.619
4min	0.755	0.601
5min	0.61	0.575

#### 4.2 異なる特徴量でのストレス状態の判別

被験者 1 の 3 分間のデータに対して特徴量の組み合わせを変更し、ストレス状態の識別を行った結果を Fig.3 に示す。特徴量として SDNN、RMSSD、SDNN/RMSSD を用いた結果は最も低い正解率の 0.65 となった。これらは似た意味をもつ特徴量であるため、これらの特徴量のみではストレス状態の判別に十分な情報が得られなかった可能性がある。瞬時心拍数の平均値、SDNN、RMSSD を用いた結果は正解率 0.76 となった。また、瞬時心拍数の平均値、分散値、RMSSD を用いた結果は正解率 0.80 となり、違いの大きな特徴量を学習に使用することで正解率が向上することが確認できた。一方、すべての特徴量を用いた識別結果は正解率 0.75 と前述した条件での識別より低い正解率となった。似た特徴量を用いることは効果が期待できないだけでなく、学習の妨げになる可能性がある。今後は瞬時心拍数の平均値、分散値、RMSSD を特徴量として使用し、さらに特徴量を追加する場合も現状の特徴量の差異が大きな情報を追加することで効果的に学習を進められるよう検討する必要がある。

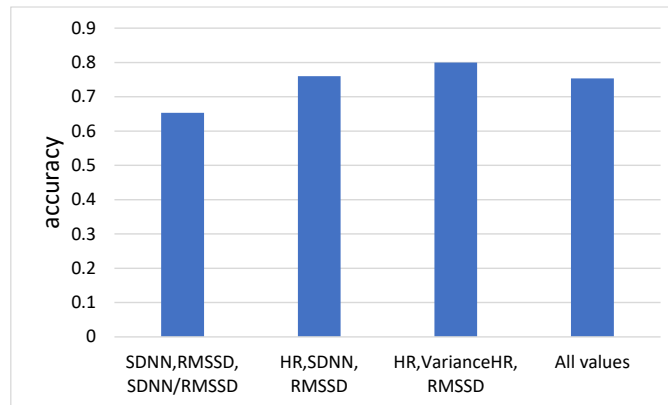


Fig.3 特徴量ごとのストレス判別結果

### 4.3 NNによるストレスレベルの推定

結果 4.1 より、ストレスの推定に適した時間幅が 3 分間のデータを使用し、解析を行った。また、特徴量としては結果 4.2 でもっとも正解率の高かった瞬時心拍数の平均値、分散値、RMSSD の組み合わせを用いた。

#### 4.3.1 被験者 1 の予測値と正解ラベルの相関関係

被験者 1 のデータセットから NN によるストレスレベルの推定を行った結果を Fig.4 に示す。被験者 1 ではストレスレベル 4 の回答結果は得られなかった。被験者 1 のストレスレベルの正解ラベルと予測値の相関係数は 0.865 となった。ストレスレベルが低いデータについては正解ラベルと同じ値に予測値が集中していることがわかる。そのため、ストレスが低い状態では高い精度で推定が可能だと考えられる。ストレスレベルが高いデータに関してはばらつきが多く見られ、ストレスレベル 2、3 のデータについてはほとんどその違いを捉えられていない。被験者 1 においてストレスレベル 2、3 のデータはそれぞれ同程度のデータ数が取得できていることから、学習用データを増やしても現状の方法ではストレスレベルが高い状態を正しく推定することは難しいと考えられる。また、そもそもストレスが一定以上の状態において被験者が正しくその心理状態を回答することが難しい可能性も考えられる。ストレスレベルが高いデータについては他の手法を検討する必要がある。

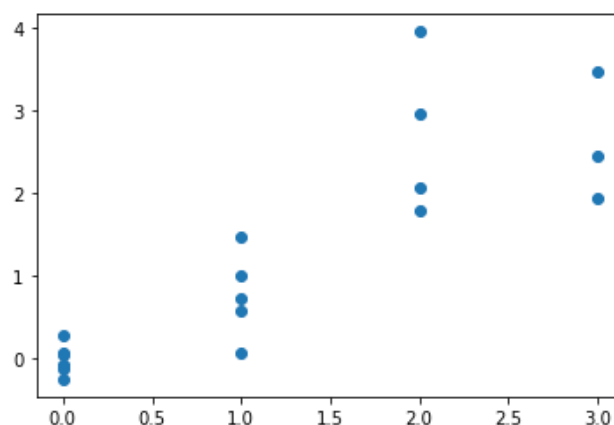


Fig.4 被験者 1 予測値と正解ラベルの相関関係

#### 4.3.2 予測値と正解ラベルの相関係数の被験者ごとの比較

Fig.5 は 10 回ずつ学習とテストを繰り返し、正解ラベルと推定結果の相関係数の平均値を比較したものである。被験者全員に正の相関関係がみられた。このことから、ウェアラブルデバイスにより取得した瞬時心拍数と POMS 短縮版の T 得点から設定したストレスレベルには関連があり、その関係性を本研究での方法において捉えることができたと考えられる。被験者 3 のように正解ラベルのストレスレベルの分布に偏りが無い被験者の相関係数が高くなる傾向が見られた。

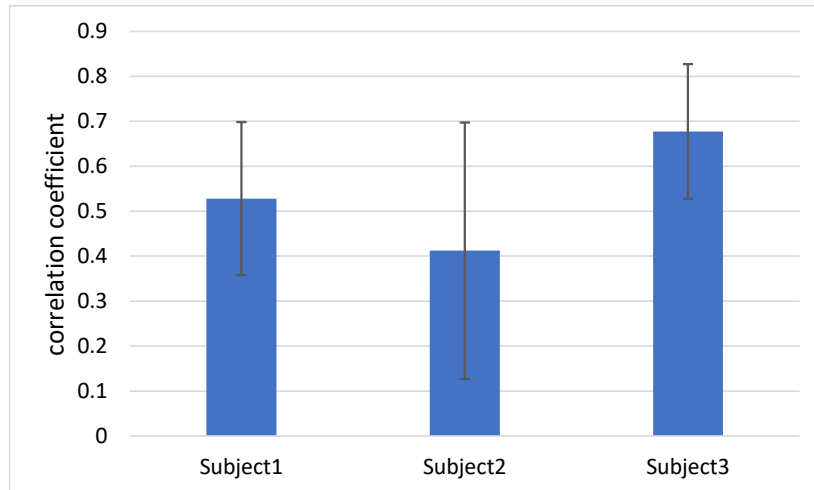


Fig.5 予測値と正解ラベルの相関係数

#### 5. 終わりに

本研究ではウェアラブルデバイスを用いて瞬時心拍数を取得、同時にストレス主観評価を行い、それらからストレス状態の識別、ストレスレベルの推定を行えるか検討した。その結果、ストレス状態の識別を行える可能性があることが分かった。瞬時心拍数の取得時間幅については2~3分間のデータが、特微量としては瞬時心拍数の平均値、分散値、RMSSD が有効であった。ストレスレベルの推定では、どの被験者にもストレスの主観評価結果と推定結果に正の相関関係が見られることから、本研究での方法において一過性ストレスを検出できる可能性がある。課題としては、現状不足している普段の生活中での強いストレスがかかった際のデータを収集することが挙げられる。

#### 参考文献

- [1] 厚生労働省, 精神疾患のデータ | 専門的な情報 | メンタルヘルス,  
<https://www.mhlw.go.jp/kokoro/speciality/data.html>
- [2] 厚生労働省,平成 30 年度 労働安全衛生調査,  
[https://www.mhlw.go.jp/toukei/list/dl/h30-46-50\\_kekka-gaiyo02.pdf](https://www.mhlw.go.jp/toukei/list/dl/h30-46-50_kekka-gaiyo02.pdf),2021/02/02
- [3] 日本健康心理学会, 2019, 「健康心理学事典」,丸善出版.
- [4] 前野創一, 「指尖容積脈波を利用したストレス評価方法の提案」,  
ライフサポート.Vol. 23 No1, 2011
- [5] 横山和仁,2005, 「POMS 短縮版 手引と事例解説」,金子書房