

BERTによるレビュー文からの場所特徴表現と パーソナリティ分析に基づくストレス推定

吉村 拓真[†] 和田 真弥^{††} 大島 裕明^{†,†††}

[†] 兵庫県立大学 大学院応用情報科学研究科 〒650-0047 兵庫県神戸市中央区港島南町 7-1-28

^{††} 株式会社 KDDI 総合研究所 〒356-8502 埼玉県ふじみ野市大原 2-1-15

^{†††} 兵庫県立大学 大学院情報科学研究科 〒651-2197 兵庫県神戸市西区学園西町 8-2-1

E-mail: [†]{aa20w514,ohshima}@ai.u-hyogo.ac.jp, ^{††}sh-wada@kddi-research.jp

あらまし 本研究では、人のパーソナリティ情報と、場所の情報から、ある人がある場所でどの程度ストレスを感じるかを推定する手法を提案する。提案手法を用いることによって、ある人がストレスを感じやすい場所やストレス解消になる場所を推薦することが可能になると考えられる。ストレス値は心拍指標である LF/HF 値を用いる。LF/HF 値を算出するために、デバイスを用いて心拍波形を測定した。しかし、デバイスによる測定は大規模に行うことが難しい。そのため、その場所で体験したストレスについて、アンケートを実施し、ストレス値を取得した。アンケートによるストレス値を訓練用のデータ、デバイスによるストレス値を検証用とテスト用のデータとして使用した。また、パーソナリティ情報として Big Five とストレス反応尺度を用いた。Big Five は 5 次元、ストレス反応尺度は 6 次元で表現し、計 11 次元のパーソナリティ情報のベクトルを取得した。場所の特徴として TripAdvisor におけるランドマークのレビュー文を用いた。レビュー文は各ランドマークに複数存在するため、レビュー文の集合をランドマークを表す特徴とした。レビュー文は BERT を用いて 768 次元のベクトルを取得した。これらのパーソナリティと場所の特徴を用いて、どのようなパーソナリティの人が、どのような場所で、どの程度のストレスを感じるかを推定する。

キーワード ストレス推定, メンタルヘルス, パーソナリティ分析

1 はじめに

近年労働や人間関係を含む日常生活において、精神的なストレスを感じることが多い、ストレス社会となっている。このような社会で、メンタルにストレス負荷のある状態のままケアを行わず放置して過ごした場合、メンタルヘルスに不調が現れ、生活に支障をきたす人も存在する。そのため、ストレスを感じた際にはメンタルヘルスケアを怠らず、ストレス解消を行う、すなわちリフレッシュすることが重要となる。

日常生活で我々がストレスを感じるシチュエーションは、「仕事や研究で成果が出せていない」や「隣人や同僚との関係がうまくいっていない」、「重大な会議を控えて緊張する」などの様に多種多様である。また、これらのシチュエーションのうち、何に対してストレスを感じるのか、というのは個人によって大きく異なり、リフレッシュの方法もそれに伴って異なる。そのため、リフレッシュ方法というのは、ストレスを感じるシチュエーションと同様に多種多様な方法が存在する。

例えば、リフレッシュの方法の例として、「山へ森林浴をしに行く」や「神社巡りをする」などがある。また、より身近な例としては「給湯室にコーヒーを入れに行く」といったものもある。これらのリフレッシュ方法には、場所を移動し、周囲の環境を変えることが共通している。そのため、人がリフレッシュをする際に、場所を移動する（周囲の環境を変える）というのは重要な要素の一つであると言える。

ある人がリフレッシュできる場所（安息地）を求めた際に、場所を選ぶ基準として、その人の過去の経験によるものがある。例えば、馴染み深い場所であれば、過去に訪れた場所の中から、お気に入りの場所を選択することができる。しかし、引越しや出張、あるいは旅行等で必ずしも馴染みの地域にいるとは限らない。そのため、ある人が安息地を求めた際に、その人がその場所へ行った時に感じるストレスを推定し、リフレッシュできる場所であるかという客観的基準が存在することが望ましい。

しかし、ある場所で感じるストレスはその場所の特徴によって変わる。例えば、人口密度の違いや、騒音による影響、周囲の建築物や自然物によって人が感じるストレスが異なる。また、山を好む人がいれば、海を好む人もいるように、ある人が感じるストレスは個人によっても異なる。そのため、ある人がある場所で感じるストレスは、その人の特徴とその場所の特徴によって決定される。

そこで本研究では、様々な人が、様々な場所でどのようにストレスを感じるかを測定し、どのようなパーソナリティの人がどのような場所で、どの程度ストレスを感じるかを推定する手法を提案する。本研究は兵庫県立大学応用情報科学研究科研究倫理委員会の研究倫理審査申請を行い承認を得た後、実施した（承認番号: UHGSAT-2021-02）。

2 節では関連研究について述べる。3 節では問題定義とデータの取得について述べる。4 節では提案手法について述べる。5 節ではアンケート結果からストレス値の取得について述べる。6 節ではレビュー文選択について述べる。7 節では結果について述べる。

て述べる。8節ではまとめについて述べる。

2 関連研究

2.1 Big Five 指標

人を心理的にパーソナライズする方法として Big Five [5], [6] という指標がある。Big Five とは人の心理は以下の 5 つに起因するとされるものである。

- Extraversion 外向性
- Agreeableness 協調性, 調和性
- Conscientiousness 勤勉性, 誠実性
- Neuroticism 神経症傾向, 情緒不安定性
- Openness 開放性

いくつかの質問項目に回答することによって Big Five を表現する手法について次のようなものがある。Big Five Scales (BFS) [21] では 5 つの下位尺度を各 12 項目, 計 60 項目の形容詞で構成された全く当てはまらない (1 点) から非常に当てはまる (7 点) の 7 段階で測定されている。短縮版 BFS (BFS-S) [18] は 5 つの下位尺度にそれぞれ 4 項目, 計 20 項目の形容詞で構成されるものである。回答は全く当てはまらない (1 点) からとてもよく当てはまる (7 点) の 7 段階で求められた。Five-Factor Personality Questionnaire 50 (FFPQ-50) [17] は 5 つの下位尺度がそれぞれ 10 項目, 計 50 項目で構成されており, 全く違う (1 点) から全くそうだ (5 点) の 5 件法で回答が求められるものである。また Gosling らの研究 [7] では 5 つの下位尺度に対して各 2 項目で計測する Ten Item Personality Inventory (TIPI) と各 1 項目で計測する Five Item Personality Inventory (FIPI) が提案されたが, TIPI の利用が推奨されている。また TIPI を日本語に訳した日本語版 TIPI (TIPI-J) [15] も存在する。本研究では人の特徴量を得るためにこの TIPI-J を用いる。

2.2 ストレス反応尺度

パーソナリティを表すために, その人のストレス特性を知る必要がある。ストレス特性を表す指標として, ストレス反応尺度というものがある。ストレス反応尺度を知るための調査を以下に示す。

- Beck Depression Inventory (BDI)
- State-Trait Anxiety Inventory (STAI)
- General Health Questionnaire (GHQ)
- Psychological Stress Response Scale (PSRS)
- Stress Response Scale-18 (SRS-18)
- Minnesota Multiphasic Personality Inventory (MMPI)

BDI [11] は 21 項目の質問に対して 4 段階の評価を行う手法であり, うつ病の判定に利用される。STAI [9] は 40 項目の質問に対して 4 段階の評価を行うことで一時的なストレス状態の状態不安とその人のストレス特性である特性不安が調査可能なものである。GHQ [4] は 60 項目の質問に対して 4 段階評価を行うことで, 6 つの尺度が調査可能なものである。また, その短縮版として GHQ30 というものがある。PSRS [16] は 53 項目

の質問について 4 段階の評価を行うことで, 情動的な項目として 4 種類, 認知行動的反応項目として 9 種類の尺度が調査可能なものである。SRS-18 [20] では 18 項目の質問項目に 4 段階評価を行うことで, 3 種類の尺度で一過的な心理変化が調査可能なものである。MMPI [8] は 550 項目に 2 段階評価を行うことで, 基礎尺度として 4 種類, 臨床尺度として 10 種類の尺度が調査可能なものである。

2.3 ストレス指標

堀ら [19] は客観的なストレス指標として用いる際の方法について議論されている。血中の Brain-derived neurotrophic factor (BDNF) 濃度については客観的なストレス指標として末梢中の BDNF の濃度は相関が低いとされた。唾液, 尿, 毛髪, 爪のようなストレスバイオマーカーについては, 試料を採取する際にサンプリング回数を等間隔で行う様に厳密に遵守しようとするれば被験者の生活リズムに大きな影響を与えてしまうことが指摘されている。心拍変動指標 (HRV) については生体試料の採取が不要である点や短時間の自律神経活動の変化からストレス状態を推定できる点が優れているとされている。また, 非接触バイタルセンシングシステムについても検討され従来の心電図と同等の心拍測定, 脈波測定ができるとされた。

Simone らの研究 [1] ではマルチモーダルな自動車運転者のストレス認識の問題について取り組まれている。心拍数 (HR), 呼吸数 (BR), 手のひらの EDA (P-EDA), 鼻周囲汗 (PER-EDA) の 4 つの信号から時間領域, 周波数領域を含む 21 の特徴量を抽出し, ストレス状態と非ストレス状態の二項対立問題として取り組まれた。

リアルタイムでストレスを取得する手法として, HRV が広く用いられている。ストレス負荷状態と休憩している状態についての HRV 指標を用いた比較をした研究 [10] では非線形からなる HRV を用いたストレスを比較している。Salahuddin ら [12] や Castaldo ら [2] の研究では HRV の内でも, 心拍の R 波と R 波の間隔である RRI (R-R Interval) に注目している。本研究では心拍変動指標の一つである LF/HF 値を用いてストレスを計測する。

2.4 自然言語処理

自然言語処理を目的としたディープラーニングモデルの一つとして Devlin ら [3] によって提案された Bidirectional Encoder Representations from Transformers (BERT) が存在する。BERT は Transformer [13] の構造が組み込まれ, 文章を文頭と文末の双方向から学習し, 文脈を読むことを実現した。また, Wikipedia などのコーパスを学習し, 日本語の汎用的な言語モデルを実現するなどの研究 [14] が行われている。

3 問題定義とデータの取得

3.1 問題定義

本節では本研究で取り組む問題について述べる。人の特徴と場所の特徴を入力として, ストレス値を推定する。ある人がある場所で感じるストレスを推定する機械学習モデルの入出力を

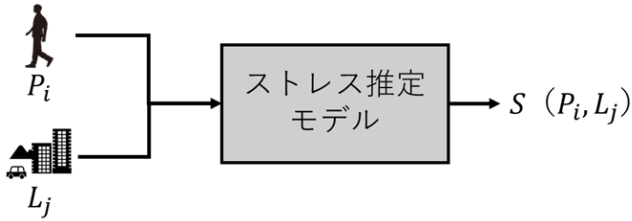


図 1 人の特徴と場所の特徴からのストレス推定

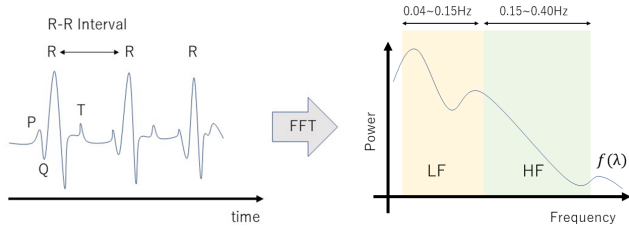


図 2 LF/HF 値の計算方法

図 1 に示す．図 1 はある人 P_i がある場所 L_j に訪れた際に感じるストレス $S(P_i, L_j)$ を推定することを意図した図である．この推定モデルでは，人の特徴と場所の特徴を入力として，ストレス値の推定を行う．そのため，ある人がある場所で感じるストレスを推定するには，ある人の特徴 P_i と場所の特徴 L_j が必要となる．本研究では人の特徴 P_i を，その人のパーソナリティで表現する．また，場所は TripAdvisor¹ に登録されている神戸市内のランドマークの内，レビュー文が 30 件以上存在するランドマークの 120 カ所を本研究における場所として定義する．ある場所の特徴 L_j はそのランドマークにおけるレビュー文の集合とし， n_j 件のレビュー文が存在する場所の特徴 L_j をレビュー文 $r_k^{L_j}$ を用いて表すと， $L_j = \{r_1^{L_j}, r_2^{L_j}, \dots, r_k^{L_j}, \dots, r_{n_j-1}^{L_j}, r_{n_j}^{L_j}\}$ となる．

ストレス値としては心拍変動指標から算出した LF/HF 値を用いる．LF/HF 値を心拍変動指標から抽出するためには心拍データを取得し，ピーク検出を行うことで R 波と R 波の間隔である R-R Interval (RRI) を抽出する必要がある．そのピーク間隔を RRI として図 2 の様に Fast Fourier Transform (FFT) でパワースペクトルを求める．その後，LF を 0.04 から 0.15 Hz，HF を 0.15 から 0.40 Hz の範囲として式 1 で LF/HF 値を計算する．

$$\frac{LF}{HF} = \frac{\int_{0.04\text{Hz}}^{0.15\text{Hz}} f(\lambda)}{\int_{0.15\text{Hz}}^{0.40\text{Hz}} f(\lambda)} \quad (1)$$

また，ある人がある場所で感じるストレス値は，個人によって取りうる値の範囲が異なる．本研究におけるストレス解消については，その個人の中で低値となる場所を提案することが望ましい．そのため，本研究ではストレス値を各個人で平均が 0 分散が 1 となるように正規化を行った相対的なストレス値を用いる．その相対的なストレス値に対して，0 をしきい値として High と Low の 2 値でラベル付けしたもので，High の場合ス

表 1 人の特徴		
項目	要素	調査アンケートの名称
p_1	外向性	TIPI-J
p_2	協調性	
p_3	勤勉性	
p_4	神経症傾向	
p_5	開放性	
p_6	一般的疾患傾向	GHQ30
p_7	身体的症状	
p_8	睡眠障害	
p_9	社会的活動障害	
p_{10}	不安と気分変動	
p_{11}	希死念慮	

表 2 TIPI-J の 10 項目の質問		
項目	質問内容	正負
1	活発で外向的だと思う	$p_{1,+}$
2	他人に不満をもち，もめごとを起こしやすいと思う	$p_{2,-}$
3	しっかりしていて，自分に厳しいと思う	$p_{3,+}$
4	心配性で，うろたえやすいと思う	$p_{4,+}$
5	新しいことが好きで，変わった考えをもつと思う	$p_{5,+}$
6	ひかえめで，おとなしいと思う	$p_{1,-}$
7	人に気をつかう，優しい人間だと思う	$p_{2,+}$
8	だらしく，うっかりしていると思う	$p_{3,-}$
9	冷静で，気分が安定していると思う	$p_{4,-}$
10	発想力に欠けた，平凡な人間だと思う	$p_{5,-}$

トレスを感じる，Low の場合ストレスを感じないとする．

本研究では入力として人の特徴と場所の特徴の 2 種類のデータを用い，ある人がある場所で感じる相対的なストレス値の推定を行う．そのために必要なデータは

- 人の特徴
- 場所の特徴
- 個人による相対的なストレス値

の 3 つとなる．

3.2 人の特徴

本研究ではある人 P_i を表す指標として，Big Five とストレス反応尺度を用いて，表 1 に示される 11 種類の要素を用いて表現する．よって，ある人 P_i は $P_i = \{p_1^i, p_2^i, \dots, p_{10}^i, p_{11}^i\}$ で表すことができる．

表 1 に示される $p_f (f = 1, 2, 3, 4, 5)$ については TIPI-J [15] を用いて調査した．また， $p_f (f = 6, 7, 8, 9, 10, 11)$ については GHQ30 を用いて調査した．

3.2.1 TIPI-J

本研究では人の特徴の要素 $p_f (f = 1, 2, 3, 4, 5)$ を TIPI-J を用いて表現する．TIPI-J は表 2 に示される 10 項目の質問に「全く違う (1 点)」から「強くそう思う (7 点)」の 7 段階の回答をする事で表 1 に示される要素を表現する．TIPI-J には表 2 に示されるように各要素 $p_f (f = 1, 2, 3, 4, 5)$ には正 $p_{f,+}$ と負 $p_{f,-}$ の項目が存在する．この $p_{f,+}$ と $p_{f,-}$ から要素 p_f の値を算出する式を式 2 に示す．

1 : TripAdvisor : <https://www.tripadvisor.jp/>

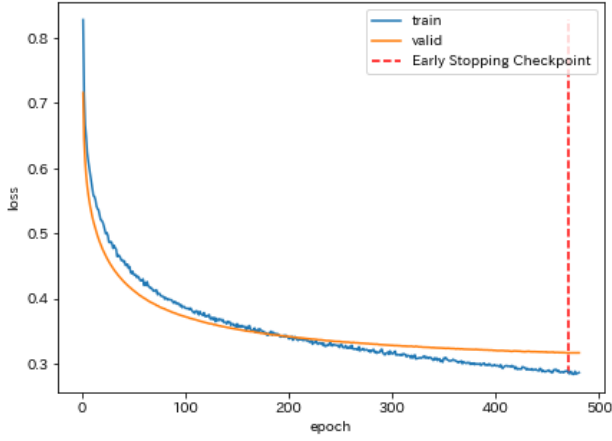


図3 TripAdvisorBERT モデルの追加学習経過図

$$p_f = \frac{p_{f,+} - p_{f,-}}{6} \quad (f = 1, 2, 3, 4, 5) \quad (2)$$

この式によって人の特徴の各要素 $p_f (f = 1, 2, 3, 4, 5)$ は -1 から 1 までの範囲で正規化された値として取得する。

3.2.2 GHQ30

本研究では人の特徴の要素 $p_f (f = 6, 7, 8, 9, 10, 11)$ を GHQ30 を用いて表現する。GHQ30 は 30 項目の質問に対して 4 段階の回答から GHQ 法による採点を行うことで 0 から 5 値の範囲で取得する。その値を $GHQ_f (f = 6, 7, 8, 9, 10, 11)$ とした場合、 $p_f (f = 6, 7, 8, 9, 10, 11)$ は式 3 を用いて求められる。

$$p_f = \frac{GHQ_f - 3}{2} \quad (f = 6, 7, 8, 9, 10, 11) \quad (3)$$

式 3 によって $p_f (f = 6, 7, 8, 9, 10, 11)$ は -1 から 1 までの範囲で正規化された値として取得する。

3.3 場所の特徴

本研究では TripAdvisor にある各場所のユーザーレビュー文をその場所の特徴として用いる。

場所の特徴ベクトルを取得するために、東北大学の Wikipedia 事前学習モデルの BERT² を用い、TripAdvisor のレビュー文で追加学習を Masked Language Model で行った。訓練用のレビューを 11,109 文、検証用のレビューを 3,703 文、テスト用のレビューを 3,704 文としてファインチューニングを行った際の学習経過図を図 3 に示す。学習条件は以下の様に設定し、損失値が 10epoch 改善されなければ Early Stopping をかけ、学習を終了した。

- batch size : 32
- 学習率 : 2e-7
- epoch : 条件によって早期終了
- 入力長 : 128

図 3 の学習は 471 epoch で終了した。テストデータによるマスク推定の正解率は 0.7355 となった。この追加学習された BERT モデルを TripAdvisorBERT として、ランドマークのレビュー文から、場所の特徴ベクトルを取得する際に使用した。

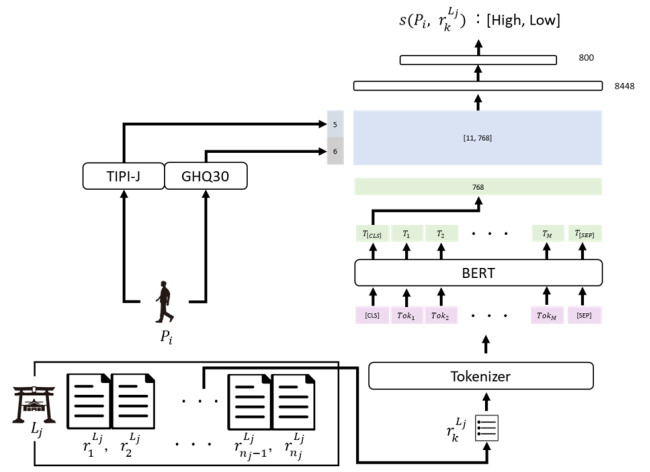


図4 人の特徴と場所の特徴からのストレス推定モデル

3.4 個人による相対的なストレス値

本研究では、ある人がある場所に訪れた際に感じるストレスを High と Low の 2 値で分類を行う。このストレスとは、その個人にとって相対的に High か Low であるかを推定するために、各個人で平均が 0 分散が 1 となるように正規化を行った相対的なストレス値を用い、0 をしきい値として High と Low でラベル付けを行った。測定地点に移動後、1 分間の安静期間を設定し、その後 5 分間心拍波形を記録した。これを 12 人の被験者で複数箇所実施し、その心拍波形から LF/HF 値を算出した。

4 人の特徴と場所の特徴からのストレス推定

人の特徴と場所の特徴から、その人がその場所へ行ったときに感じる相対的なストレスを推定する手法を提案する。人の特徴と場所の特徴を入力としてストレスを推定するモデルを図 4 に示す。

図 4 は人 P_i がランドマーク L_j に訪れた際に感じるストレス値を推定することを想定した図である。人 P_i は TIPI-J と GHQ30 を用いて $P_i = \{p_1^i, p_2^i, \dots, p_{10}^i, p_{11}^i\}$ の 11 次元で表される。また、ランドマーク L_j は n_j 個のレビュー文 r^{L_j} の集合である。そのため、ランドマーク L_j は $L_j = \{r_1^{L_j}, r_2^{L_j}, \dots, r_{n_j-1}^{L_j}, r_{n_j}^{L_j}\}$ と表すことができる。このレビュー文の内、任意のレビュー文 $r_k^{L_j}$ を選択し、Tokenizer を用いて、トークン化を行う。トークン化されたレビュー文 $r_k^{L_j}$ を 3.3 節で追加学習した BERT に入力し、768 次元の CLS ベクトルを取得する。この 768 次元のレビュー文 $r_k^{L_j}$ の CLS ベクトルと 11 次元のパーソナリティ情報 P_i のそれぞれの積を求め、 $[11, 768]$ のマトリクスを計算し、平滑化を行い、8,448 次元のベクトルとする。そして、 P_i と $r_k^{L_j}$ から推定されたストレス値 $s(P_i, r_k^{L_j})$ を推定する。しかし、これは人 P_i とランドマークの内の 1 つの要素であるレビュー文 $r_k^{L_j}$ からのストレス値 $s(P_i, r_k^{L_j})$ の推定であるため、人の特徴と場所の特徴からのストレス推定ではない。

ここで本研究では人 P_i とランドマーク L_j からのストレス推定手法として多数決を行う。ある人 P_i があるランドマーク L_j で感じるストレス $S(P_i, L_j)$ を式 4 のようにレビュー文 $r_k^{L_j}$ ご

2 : BERT 事前学習モデル : <https://github.com/cl-tohoku/bert-japanese>

とに求めたストレス値 $s(P_i, r_k^{L_j})$ の集合 $S(P_i, L_j)$ としてみなし、 $s(P_i, r_k^{L_j})$ で多数決を行うことで、ある人 P_i があるランドマーク L_j で感じるストレス $S(P_i, L_j)$ を求める。

$$S(P_i, L_j) = \{s(P_i, r_k^{L_j}) | k = 1, 2, \dots, n_j - 1, n_j\} \quad (4)$$

この多数決において、 $s(P_i, L_j)$ で High が半数以上である場合に $S(P_i, L_j)$ を High とし、 $s(P_i, L_j)$ で High が半数未満である場合に $S(P_i, L_j)$ を Low とした。

このストレス推定モデルを学習するにあたって、問題点が2つ存在する。一つ目に、このモデルの構築には多くの訓練データが必要となる。デバイスを用いた測定では大規模なデータ収集が困難であるため、アンケート結果からストレス値を取得することについて取り組む。二つ目に、ランドマーク L_j は n_j 個のレビュー文 r^{L_j} の集合であるため、その内のいくつかは、その人がその場所へ行ったときに感じる相対的なストレスを推定するために重要ではないレビュー文を含むことで、場所の表現に冗長性があることが考えられる。そのため、ストレスを推定するにあたって重要となるレビュー文の選択について取り組む。

5 アンケート結果からストレス値の取得

本研究における課題として、ストレス値の大規模なデータの取得が困難であることが挙げられる。本研究でのストレス測定の実験は、被験者が心拍測定デバイスを装着し、測定地点へ移動してその場で1分の安静期間を経て5分間の測定を行う。この測定方法は被験者の移動を伴うことから、測定に時間がかかることが問題点としてある。また、デバイスには数的制限があるため、測定を並列に行うにも上限がある。そのため、心拍変動指標を用いてストレス値を大規模に取得することは困難である。

この代替手段として、その場所で感じたストレスを調査するアンケートから、その場所で感じたストレス値を推定する手法を提案する。その場所で感じたストレスを調査するアンケートはストレスを感じる、あるいはリフレッシュできるとされる状況から計20項目の質問を作成した。その20項目の中には冗長な質問項目が含まれていたため、実際のストレスと相関のある重要な項目を抽出した。そのアンケート結果からストレス値を取得した。

5.1 Landmark Stress Questionnaire (LSQ) の作成

アンケートは場所に関するストレス調査を行うために下記の5つの要素を想定した。

- 人口密度
- 騒音
- リラックス
- 非日常性
- エンターテインメント性

これらの各要素に対してそれぞれ4問、計20項目の質問を設定し、「全くそう思わない」から「強くそう思う」の7段階で回答することで、その人がその場所で感じるストレス値を表現し、

表3 LSQの20項目の質問

項目	質問文
1	人の多さで不快に感じる場所であったか？
2	人が多くて楽しいと感じる場所であったか？
3	周囲の人の密度が気になる場所であったか？
4	人が少なく寂しいと感じる場所であったか？
5	騒音が気になるような場所であったか？
6	にぎやかで楽しい場所であったか？
7	音がせず、寂しいと感じる場所であったか？
8	静かで心地よい場所であったか？
9	気が散るような場所であったか？
10	落ち着ける場所であったか？
11	憂鬱な気分になる場所であったか？
12	和む場所であったか？
13	仕事や学業等を思い出すような場所であったか？
14	意外性のある場所であったか？
15	ありふれた場所であったか？
16	非日常的な場所であったか？
17	つまらない場所であったか？
18	有意義な時間を過ごせる場所であったか？
19	殺風景な場所であったか？
20	好みの場所であったか？

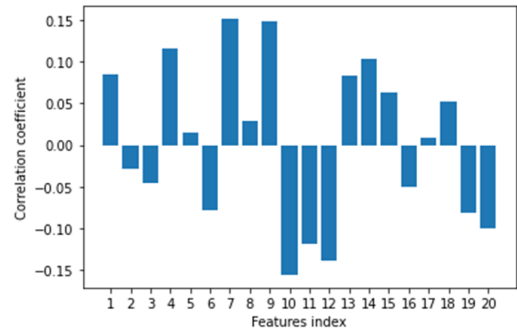


図5 LSQ 質問項目とストレスの相関

実測したものの代用としてデータを用いることができるか検証する。このアンケートの20項目を表3に示す。

表3に示される質問項目の中には、冗長な質問や、実際はストレスと相関の無い質問項目が存在することが考えられる。そのため、実際のストレスと相関があるかを確認するために、被験者12名に対して、計120カ所のランドマークに対し、神戸市内のランドマークに訪れてもらい、デバイスを用いたストレス推定と、その各ランドマークごとに表3のアンケートを実施した。

5.2 LSQの重要な質問項目の抽出

実際に測定したストレス値と表3で実施したアンケートの結果との相関を示した結果を図5に示す。

図5から、質問項目1から20までのすべてにおいて、相関係数の絶対値が0.15より大きい値を持つものがないため、表3とストレスとの相関は小さいことが読み取れる。

表3の質問項目はその人個人の主観によるものが大きく、パーソナリティによって回答に影響が出ると考えた。3.2.1節

表 4 LSQ の選択された質問項目とパーソナリティの組み合わせ

LSQ の項目	パーソナリティ	相関係数
4	p_{10}	0.361
5	p_2	0.338
6	p_7	-0.367
7	p_{10}	0.413
	p_{11}	0.315
9	p_2	0.302
	p_6	0.301
14	p_{10}	0.371
	p_{11}	0.325
19	p_{10}	0.337

でのパーソナリティ、Big Five と、3.2.2 節でのパーソナリティ、ストレス反応尺度の計 11 要素と表 3 のアンケートの回答結果の積をもとめ、実測したストレス値との相関を確認したものを図 6 に示す。図 6 は正の相関が大きいほど濃い青に、負の相関が大きいほど濃い赤に、相関が無い場合は白に近い色としている。

図 6 を見ると、図 5 で得られた相関係数より、大きな値を持つものが出てきていることがわかる。この内、相関係数の絶対値が 0.300 以上のものを対象とし、アンケート項目からストレス値を推定するために重要な質問項目として、採用した。そのアンケート項目とパーソナリティの組み合わせを表 4 に示す。

表 4 の抽出されたアンケート項目を LSQ の重要な 7 項目、LSQ7 とし、これらとパーソナリティの組み合わせをアンケートからストレス値を取得するための特徴として使用した。

5.3 LSQ7 の結果からのストレス値の推定

ここにアンケートからのストレス推定の手順を示す。アンケート結果からのストレス値の推定には、表 4 に示される質問項目とパーソナリティの積である、計 10 種類の値を説明変数とする。この 10 種類の説明変数を線形回帰モデルを用いてストレス値の推定を行う。このモデルの作成には 3.4 節で取得した LF/HF 値からのストレス値を用いた。被験者 12 名のデータを用い、Leave-one-person-out 法で推定精度を決定係数 R^2 で求めた。平均決定係数 R_{ave}^2 は 0.5309 となり、このモデルをアンケート結果からのストレス推定に使用した。データの収集は 132 名にそれぞれ複数カ所についてアンケートを実施し、計 1051 個のデータを取得した。

6 レビュー文選択

レビュー文選択は、ある人 P_i があるランドマーク L_j で感じるストレスを求める際に、ランドマーク L_j を推定する際に重要となるレビュー文 $r_k^{L_j}$ を選択する。レビュー文の中には推定を助長するものもあれば、阻害するものもある。そのため、レビュー文で場所の特徴を表現し、ストレスを推定するにあたって、重要となるレビュー文を選択する手法を提案する。

レビュー文選択は、図 4 に示される基本モデルのストレス推定におけるレビュー文 $r_k^{L_j}$ ごとの正解率 $a(r_k^{L_j})$ を、そのレビュー文の重要度を表す指標として採用する。あるレビュー文

表 5 学習に用いたデータの数

使用用途	データ数
訓練データ	431,699
検証データ	56,275
テストデータ	43,340

$r_{k_1}^{L_j}$ は、そのランドマーク L_j に訪れた人の数だけストレス値 $s(P_i, r_{k_1}^{L_j})$ が存在する。レビュー文 $r_k^{L_j}$ ごとに正解率 $a(r_k^{L_j})$ を求め、人の特徴と場所の特徴からストレスを推定するにあたって、あるレビュー文が重要であるかの判断を行う。

ストレス推定で用いられる訓練データに含まれるレビュー 8,629 文を訓練：検証：テストが 8：1：1 となるようにデータを分割し、レビュー文 $r_k^{L_j}$ からレビュー重要度 $a(r_k^{L_j})$ の推定を図 7 のモデルを用いて学習を行った。学習条件は以下の様に設定し、損失値が 10epoch 改善されなければ Early Stopping をかけ、学習を終了した。

- batch size：32
- 学習率：2e-7
- epoch：条件によって早期終了
- 入力長：128

学習は 117epoch で終了し、テストデータで推定を行った結果、決定係数が 0.9397 となった。この結果から、レビュー文の重要度を推定するにあたって、十分な決定係数であると判断し、このモデルをレビュー重要度推定モデルとして用いた。このレビュー重要度 $a(r_k^{L_j})$ をランドマーク L_j ごとにレビュー文 $r_k^{L_j}$ をランキングし、上位 n 件をそのランドマークを表現するレビュー文として採用し、図 4 で人 P_i とランドマークの内の 1 つの要素であるレビュー文 $r_k^{L_j}$ からのストレス値 $s(P_i, r_k^{L_j})$ の推定を行い、多数決である人 P_i があるランドマーク L_j で感じるストレス $S(P_i, L_j)$ を求める。

7 結果

人の特徴と場所の特徴からストレスの推定を行った結果について述べる。各レビュー文ごとにストレスを推定し、その後多数決を取る手法で表 5 に示されるデータを用いて推定を行う。表 5 に示されるデータ数は人とレビュー文の組み合わせ $(P_i, r_k^{L_j})$ の数である。訓練データは 5.3 節で述べたアンケート結果から取得したストレス値によって正解ラベルが付与されているデータである。検証データとテストデータは心拍波形から取得したデータを正解ラベルとし、人数が 5：5 となるように分割した。

まず、このデータを用いてレビュー文の選択を行わず、学習条件を以下の様に設定し、損失値が 10epoch 改善されなければ Early Stopping をかけ、学習を終了した。

- batch size：32
- 学習率：2e-8
- epoch：条件によって早期終了
- 入力長：128

推定した結果、正解率は 0.7007 となった。この推定された結

		LSQの質問項目																			
パーソナリティ		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
	p_1	-0.015	-0.034	-0.057	0.140	-0.086	-0.094	0.145	0.069	-0.032	0.063	-0.045	-0.042	-0.001	0.217	-0.090	-0.040	0.055	0.001	0.042	0.030
	p_2	0.211	0.230	0.141	0.136	0.338	0.250	0.026	-0.050	0.302	-0.050	0.124	0.053	0.178	0.091	0.221	0.114	0.083	0.153	0.049	0.034
	p_3	0.000	0.155	0.053	0.020	0.049	0.162	-0.055	0.006	0.010	0.095	0.042	0.177	-0.146	0.128	-0.049	0.152	-0.022	0.072	0.052	0.121
	p_4	0.039	0.099	0.041	0.086	0.073	0.091	0.021	-0.079	0.043	-0.006	-0.009	0.035	-0.032	0.097	0.049	0.114	0.004	0.092	0.058	0.040
	p_5	-0.130	-0.176	-0.141	-0.150	-0.217	-0.190	-0.062	-0.037	-0.087	-0.082	-0.103	-0.192	-0.078	-0.083	-0.180	-0.093	-0.154	-0.064	-0.137	-0.063
	p_6	-0.053	-0.077	-0.145	0.171	-0.138	-0.183	0.184	0.088	-0.024	0.007	-0.069	-0.072	-0.067	0.301	-0.130	-0.076	0.037	-0.070	0.001	-0.013
	p_7	-0.167	-0.210	-0.226	0.140	-0.203	-0.367	0.166	0.092	-0.068	-0.003	-0.092	-0.089	-0.093	0.161	-0.189	-0.204	0.056	-0.230	0.011	-0.101
	p_8	0.017	-0.005	-0.033	0.173	-0.057	-0.041	0.232	0.139	0.020	0.019	-0.074	0.050	-0.075	0.232	-0.043	0.083	-0.018	0.063	-0.019	0.019
	p_9	-0.106	-0.116	-0.157	0.028	-0.159	-0.164	0.073	-0.063	-0.073	-0.053	-0.105	-0.117	-0.161	0.095	-0.148	-0.071	-0.081	-0.110	-0.026	-0.124
	p_{10}	0.231	0.143	0.120	0.361	0.106	0.086	0.413	0.246	0.099	0.145	0.052	0.228	0.108	0.371	0.138	0.112	0.193	0.157	0.337	0.196
	p_{11}	0.066	0.111	-0.005	0.252	-0.052	0.060	0.315	0.184	0.042	0.068	-0.034	0.127	-0.027	0.325	-0.048	0.185	0.078	0.259	0.126	0.230

図 6 LSQ 質問項目とパーソナリティの積とストレス値の相関

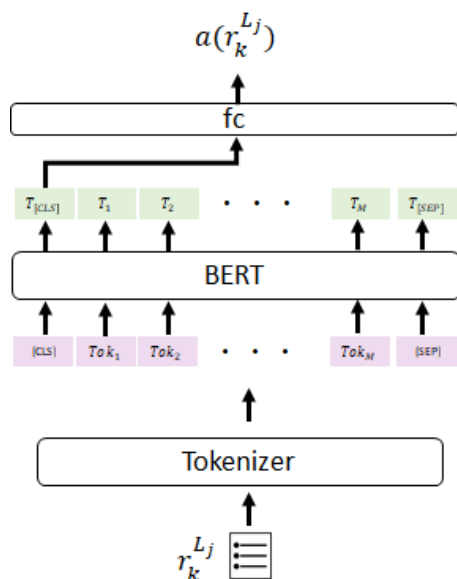


図 7 レビュー重要度推定モデル

果に対しランドマークごとに多数決を取った結果、正解率が 0.6774 となった。レビュー文ごとに正誤の判定を行った結果の正解率 0.7007 と、ランドマーク内でレビュー文の多数決を行い、正誤の判定を行った結果の正解率 0.6774 の差は、各ランドマークにおけるレビュー文の数の差が影響していることが考えられる。

そのため、各ランドマークのレビュー文の数を揃えるために、レビュー文選択手法を用いる。レビュー文の選択はレビュー文ごとに重要度の推定を行い、そのランキングの上位 n 件のレビュー文を選択する。今回、 n は 1 から 29 までの奇数の値を用いて、レビュー文を何件選択すると最も正解率が良くなるかを比較し検証した。学習条件は以下の様に設定し、損失値が 10epoch 改善されなければ Early Stopping をかけ、学習を終了した。

- batch size : 32
- 学習率 : 2e-6
- epoch : 条件によって早期終了
- 入力長 : 128

図 8 より、レビュー文の選択数 n が 13 の時に最も良い正解率の 0.8642 となった。ランドマークごとのレビュー数の内、特

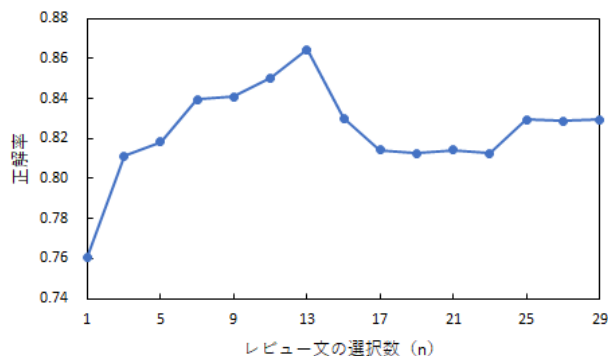


図 8 レビュー文選択を用いたランドマークごとのストレス推定結果

表 6 テストデータの被験者に推薦された場所

人	中華街 (南京町)	三宮センター街	生田神社	布引の滝
P_1	○	○	○	×
P_2	○	×	○	○
P_3	×	×	○	○
P_4	○	×	×	○

にレビュー数が 30 件に近いランドマークにおいては、25 件や 29 件を選択しても、実際のレビューの大半が採用されるため、レビュー選択の影響が小さいことが考えられる。また、そのレビュー数のおよそ半数である 15 件から、正解率が上昇していることが確認できる。そのため、レビューの選択数が 13 件で正解率が最高値となったのは、レビュー数の少ないランドマークにおいても、レビュー選択の影響が出たことが考えられる。しかし、13 件より少ない選択数ではレビューが減少するごとに正解率も減少している。これは、人の特徴によって正解しやすいレビュー文が変化するため、その場所の特徴の多様性を維持出来なくなったためだと考えられる。

また、このレビュー選択数 13 件の推定結果において、テストデータの被験者に推薦された場所の一例を表 6 に示す。 P_1 は外向性 p_1 が高い、 P_2 は外向性 p_1 が低い、 P_3 は協調性 p_2 が低い、 P_4 は一般的疾患傾向 p_6 が高い例として、推薦された場所を○、推薦されなかった場所を×で表している。表 6 を見ると、パーソナリティの違いによって、推薦される場所が異なっていることから、人によってある場所で感じるストレスの違いを考慮できていると言える。

8 ま と め

本研究では、人の特徴と場所の特徴から、その人がその場所で感じるストレスを推定する手法について取り組んだ。人の特徴としては、Big Five 指標とストレス反応尺度を用いた。Big Five 指標は TIPI-J を用いて取得し、ストレス反応尺度は GHQ30 を用いて取得した。場所の特徴は、TripAdvisor にあるレビュー文を BERT を用いてベクトル化することでその場所を表現した。本研究においてストレス値としては2種類存在する。一つは心拍波形から取得したストレス値、もう一つはアンケート結果から取得したストレス値である。心拍波形から取得したストレス値は、実際にいくつかのランドマークに訪問し、心拍センサを用いて取得した。アンケート結果からのストレス値は、ストレス値が取得できるようなアンケート LSQ を作成し、LSQ の結果からストレス値を取得した。これらのデータを用いて、アンケートで取得したストレス値を訓練データとして使用し学習を行い、心拍波形から取得したストレス値を検証のデータとテストデータとして正解率を算出した。その結果、多数決での正解率が 0.6774 となった。また、レビュー文選択手法ではランドマークごとにレビュー文を n 件選択し、学習を行い比較を行った結果、13 件で最も正解率が高く 0.8642 となった。また、テストデータの被験者でリフレッシュ場所の推薦を行ったところ、パーソナリティの違いによって、推薦される場所が異なったため、人の特徴によってある場所で感じるストレスの違いを考慮できていることがわかった。今後は、今回扱わなかった動的に変化する場所の特徴、例えば天気や季節、時間帯等を用いて人の特徴と場所の特徴からのストレス推定を行う。

謝 辞

本研究は JSPS 科研費 JP21H03775, JP21H03774, JP21H03554, JP18H03244, ならびに、2021 年度国立情報学研究所公募型共同研究 (21S1001) の助成を受けたものです。ここに記して謝意を表します。

文 献

- [1] Simone Bianco, Paolo Napoletano, and Raimondo Schettini. Multimodal Car Driver Stress Recognition. In *PCTH*, pp. 302–307, 2019.
- [2] Rossana Castaldo, Paolo Melillo, Umberto Bracale, Mariangela Caserta, Maria Triassi, and Leandro Pecchia. Acute Mental Stress Assessment Via Short Term HRV Analysis in Healthy Adults: A Systematic Review with Meta-Analysis. *Biomedical Signal Processing and Control*, Vol. 18, pp. 370–377, 2015.
- [3] Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. In *NAACL-HLT*, pp. 4171–4186, 2019.
- [4] David P. Goldberg and Valerie F. Hillier. A Scaled Version of the General Health Questionnaire. *Psychological medicine*, Vol. 9, No. 1, pp. 139–145, 1979.
- [5] Lewis R. Goldberg. An Alternative “Description of Personality”: The Big-Five Factor Structure. *Journal of Personality and Social Psychology*, Vol. 59, No. 6, pp. 12–16, 1990.
- [6] Lewis R. Goldberg. The Development of Markers for the Big-Five Factor Structure. *Psychological Assessment*, Vol. 4, No. 1, pp. 26–42, 1992.
- [7] Samuel D. Gosling, Peter J. Rentfrow, and William B. Swann Jr. A Very Brief Measure of the Big-Five Personality Domains. *Journal of Research in Personality*, Vol. 37, No. 6, pp. 504–528, 2003.
- [8] Jordan B. Hiller, Robert Rosenthal, Robert F. Bornstein, David T.R. Berry, and Sherrie Brunell-Neuleib. A Comparative Meta-Analysis of Rorschach and MMPI Validity. *Psychological Assessment*, Vol. 11, No. 3, p. 278, 1999.
- [9] Theresa M. Marteau and Hilary Bekker. The Development of a Six-Item Short-Form of the State Scale of the Spielberger State-Trait Anxiety Inventory (STAI). *British Journal of Clinical Psychology*, Vol. 31, No. 3, pp. 301–306, 1992.
- [10] Paolo Melillo, Marcello Bracale, and Leandro Pecchia. Non-linear Heart Rate Variability Features for Real-Life Stress Detection. Case Study: Students under Stress Due to University Examination. *BioMedical Engineering Online*, Vol. 10, No. 1, pp. 1–13, 2011.
- [11] Paul Richter, Joachim Werner, Andres Heerlein, Alfred Kraus, and Heinrich Sauer. On the Validity of the Beck Depression Inventory. *Psychopathology*, Vol. 31, No. 3, pp. 160–168, 1998.
- [12] Lizawati Salahuddin, Jaegeol Cho, Myeong Gi Jeong, and Desok Kim. Ultra Short Term Analysis of Heart Rate Variability for Monitoring Mental Stress in Mobile Settings. In *EMBS*, pp. 4656–4659, 2007.
- [13] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Ukasz Kaiser, and Illia Polosukhin. Attention is All You Need. In *NIPS*, pp. 6000–6010, 2017.
- [14] 柴田知秀, 河原大輔, 黒橋禎夫. BERT による日本語構文解析の精度向上. 言語処理学会, pp. 205–208, 2019.
- [15] 小塩真司, 阿部晋吾. 日本語版 Ten Item Personality Inventory (TIPI-J) 作成の試み. 日本パーソナリティ心理学会 パーソナリティ研究, Vol. 21, No. 1, pp. 40–52, 2012.
- [16] 新名理恵, 坂田成輝, 矢富直美, 本間昭. 心理的ストレス反応尺度の開発. 一般社団法人 日本心身医学会 心身医学, Vol. 30, No. 1, pp. 29–38, 1990.
- [17] 藤島寛, 山田尚子, 辻平治郎. 5 因子性格検査短縮版 (FFPQ-50) の作成. 日本パーソナリティ心理学会 パーソナリティ研究, Vol. 13, No. 2, pp. 231–241, 2005.
- [18] 内田照久. 音声の発話速度が話者の性格印象に与える影響. 公益社団法人 日本心理学会 心理学研究, Vol. 73, No. 2, pp. 131–139, 2002.
- [19] 堀輝, 香月あすか, 菅健太郎, 吉村玲児. 客観的なストレス評価方法について. 日本職業・災害医学会会誌, Vol. 66, No. 5, pp. 330–334, 2018.
- [20] 鈴木伸一, 嶋田洋徳, 三浦正江, 片柳弘司, 右馬埜力也, 坂野雄二. 新しい心理的ストレス反応尺度 (SRS-18) の開発と信頼性・妥当性の検討. 日本行動医学会 行動医学研究, Vol. 4, No. 1, pp. 22–29, 1997.
- [21] 和田さゆり. 性格特性用語を用いた Big Five 尺度の作成. 公益社団法人 日本心理学会 心理学研究, Vol. 67, No. 1, pp. 61–67, 1996.