

労働者の精神衛生モニタリングに向けたデータ収集と分析

前川 哲志¹ 水本 旭洋² 諏訪 博彦^{1,3} 島津 明人⁴ 安本 慶一¹

概要: 現在, 我が国の労働生産性は先進7カ国(G7)で最下位となっており, 健康を維持したまま生産性を向上させる働き方改革が急務とされている。働き方改革においては, 労働者の好調・不調の程度を示す「Work Attitude」(ワーク・エンゲージメントやリカバリーなど)に基づく健康管理が必須であり, 日常的に Work Attitude を計測することが求められている。一方で, 従来の Work Attitude の評価手法は, 1年に1回程度のアンケートによるものであり, モニタリングに適したものではない。そこで, 本研究では, 主観的アンケートに基づいて散発的にのみ計測されていた Work Attitude を, マルチモーダル情報を用いて持続的に計測・記録する「Work Attitude PLR (Personal Life Record) 収集基盤」を構築する。本発表では, 収集基盤の構築に向けて行ったデータ収集実験とその分析について述べる。

キーワード: 働き方改革, ワーク・エンゲイジメント, マルチモーダル情報, 感情推定

1. はじめに

現在, 我が国では仕事や職業生活に関することで強いストレスを感じている労働者の割合がおよそ60%に及んでおり [1], 精神に支障をきたす労働者が増加傾向にある [2]。その精神障害発症に影響を与えた業務の負荷要因調査では, 男女共通して恒常的な長時間労働が要因として多く, 精神に支障をきたす原因の多くが長時間労働である [3]。長時間労働の問題が深刻であるにも関わらず, 労働生産性は低く [4], 長時間労働のわりに労働生産性が低いのが現状である。この問題の解決には, 政府が推進している働き方改革のように, 労働者一人ひとりが精神的・身体的な健康を維持したまま生産性を向上させることが重要である。そのため, 労働者自身が自分の働き方を客観的に振り返ることが重要であると考え。労働者自身が自分のストレス状態や, 仕事に対してどれだけやりがいを持って取り組んでいるかといった情報を日常的に取得することによって, より効率的な働き方の発見やワークスタイルの見直しにつながると考える。

したがって, 労働者の精神衛生の日常的なモニタリングを行い, 労働者の仕事に対する状態を示す「Work Attitude」

(ストレスやワーク・エンゲイジメントなど, 労働者の好調・不調を評価する様々な指標をまとめたものと定義する) を労働者自身が取得することが必要であると考えられる。我が国における, 労働者の精神衛生をモニタリングする既存方法として, 職業性ストレス簡易調査票を用いたアンケート [5] が挙げられる。これは労働安全衛生法に基づいて, 労働者数50人以上の全ての事業場に対して実施が義務付けられている。しかしながら実施頻度は1年に1回程度であり, 日常的なモニタリングとしては用いられていない。また, 多数の質問項目に答えなければいけないため, 毎日の計測は負担が大きく, 継続的なモニタリングには適していない。

日常的なモニタリングを可能とするには, 誰もが所持していると考えられるスマートフォンの利用が考えられる。スマートフォンの普及に伴い, モバイルデバイスを使用してユーザの精神状態を監視し, セルフケアを促進することへの関心が高まっており [6], 本研究でもスマートフォンを用いて労働者の精神状態を推定する。

2. 関連研究

スマートデバイスを用いて心理尺度を推定しようとする研究は, 最近注目されている。雨森らは, スマートデバイスを用いた学生のQOL簡易評価に関する研究 [7] を行っている。生体データや活動データを特徴量とし, QOLの推定モデルを構築している。Natashaらは, スマートデバイスを用いた学生の幸福などを予測する研究 [8] を行っている。生体データや活動データを特徴量とし, 健康・エネ

¹ 奈良先端科学技術大学院大学

Nara Institute of Science and Technology

² 大阪大学

Osaka University

³ 理化学研究所 革新知能統合研究センター

RIKEN Center for Advanced Intelligence Project

⁴ 慶應義塾大学

Keio University

ルギー・ストレスなどの幸福さに関わるような5つの尺度を組み合わせた複合的な心理尺度を推定するモデルを構築している。Sanoらは、ウェアラブルセンサとスマートフォンによるデータ収集とアンケートによるストレス、精神的健康状態に関する調査を実施し、客観的データを使用してアカデミックパフォーマンス（GPA）やピッツバーグ睡眠指標（PSQI）の値を67～92%の分類精度の範囲で推定している [9]。Boukhechpaらは、スマートフォンによる自動客観センサデータの収集と、アンケートによる社会的不安およびうつ病の病状に関する調査を実施し、データと心理的指標との間の多くの重要な相関関係を示している [10]。Fukazawaらは、スマートフォンのセンサーログとアプリケーション履歴データと不安を表すSTAIスコアから、日常生活における無意識の不安の変化を予測しようと試み、74.2%のF値を示す性能の推定モデルを構築している [11]。これらの先行研究では、主観的評価である心理指標を、ウェアラブルデバイスまたは、スマートフォンによって収集される客観的なデータから一定の水準で推定することに成功している。

しかしながら、スマートウォッチや特殊なウェアラブルデバイスなど、ユーザが特定のデバイスを所持している必要があるため、ユーザの参加コストが高い。また、殆どの研究がストレスやうつ病などのネガティブな指標を対象としており、労働者の好調さをあらわすポジティブな指標は対象とされていない。ストレスがあったとしても、仕事に対する意欲やリカバリーが高い場合は問題ないことが指摘されており、「Work Attitude」を計測するためには、ポジティブ・ネガティブ両面の指標を対象とする必要がある。

我々は、労働者の所持率が高いスマートフォンのみを用いて、アンケートレスかつ日常的にポジティブ・ネガティブ両面の心理尺度を含んだ「Work Attitude」を推定することを目的とする。

3. 提案手法

2章で述べた目的を達成するために、本研究では、労働者の好調・不調を日単位で記録する Work Attitude PLR 収集基盤を提案する。

3.1 提案システム

図1に Work Attitude PLR 収集基盤の構成を示す。システムは、データを収集するスマートフォンとそれを蓄積・処理するサーバで構成される。スマートフォンは、動画により顔の表情、声色、コメントが収録されるとともに、一日の活動量など行動データを収集し、サーバに送信する。サーバは、労働者のスマートフォンからデータが送られてくると、特徴量抽出プログラムによりデータ群から特徴量を抽出する。次に得られた特徴量を Work Attitude 推定モデルに入力することで、Work Attitude の推定値を出力す

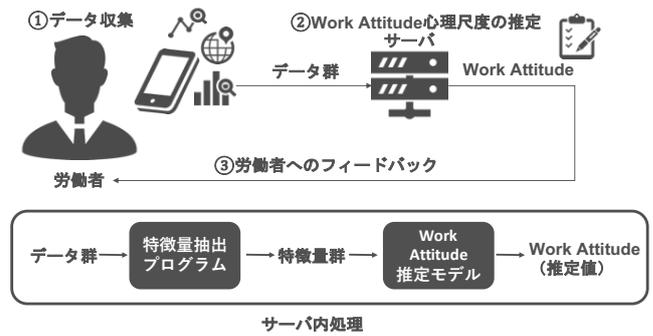


図1 Work Attitude PLR 収集基盤の構成

る。推定モデルが推定した Work Attitude は、PLR（パーソナルライフレコード）として DB に記録される。Work Attitude は労働者にフィードバックされ、これにより労働者は客観的に働き方を振り返ることが可能となる。

Work Attitude 推定モデルは、労働者の好調・不調の程度を表す複合尺度「Work Attitude 評価尺度」を開発し、事前にデータ収集し、機械学習を用いて学習することで構築する。

3.2 Work Attitude 評価尺度

労働者の精神衛生を評価するために様々な尺度が利用されているが、その多くはいくつかの下位尺度を含んだ単一の心理尺度である。また、これらの多くの尺度は、年1回の職業性ストレス指標のように、長期的な間隔で心理尺度を評価するために開発されており、項目数が多く、日常的な評価には適していない。一方で近年、日常的な心理尺度を評価することに着目した研究がいくつか行われている。日常的に評価を行う場合、項目数が多くなるほど、評価にかかる負担が大きくなるため、少ない項目数で評価が行える尺度が開発されている。本研究では、日常的に評価することを目的に開発された尺度を基に、複合的な尺度を評価可能な Work Attitude 評価尺度を開発する。

本研究では、表1に示した、ワーク・エンゲイジメント、リカバリー経験、リカバリー状態、睡眠状態、感情といった5つの尺度を、Work Attitude 評価尺度のパイロットスケールとする。これらの指標は、日常的な評価が可能なものであり、ポジティブ・ネガティブ両面の指標を含んでいる。ワーク・エンゲイジメントは、仕事に積極的に向かい活力を得ている状態を評価する心理尺度である。本研究では、活力、熱意、没頭という3つの下位尺度を持つユトレヒト・ワーク・エンゲイジメント (UWES) [17][18][19][20][21] について、3項目で評価可能な UWES-3[12] の日本語版を用いる。この尺度は、労働者の労働生産性にも関わる重要な尺度である。

リカバリー経験 (Recovery Experience Measures) [13][22] は、ストレス負荷により高まったストレスレベルを回復するための活動が余暇に行えたかを評価する尺度である。こ

表 1 Work Attitude 尺度に使用する心理尺度

評価項目	尺度	概要	項目
ワークエンゲイジメント (WE)[12]	UWES	仕事から活力を得ているか	WE1:活力 WE2:熱意 WE3:没頭
リカバリー経験 (RE)[13]	Recovery Experience Measures	余暇で回復する活動を行えたか	RE1:仕事のことを忘れる RE2:リラックス RE3:知的探求 RE4:自己決定
リカバリー状態 (RS)[14]	State of being recovered in the morning	朝回復している状態であるか	RS1:身体的回復 RS2:精神的回復 RS3:睡眠の質
睡眠状態 [15]	Pittsburgh Sleep Quality Index	睡眠の長さや質	
感情 [16]	UWIST	気分	

の尺度は、Psychological detachment（仕事からの離脱性）、Relaxation（リラクゼーション）、Mastery（仕事以外での挑戦）、Control（余暇の制御）という4つの下位尺度を評価する尺度である。各下位尺度について4項目、合計16項目に回答する必要があるが、本研究では、複合的な尺度を開発するため、各尺度に対して、因子負荷量が最も高い項目を1項目ずつ抽出し、下位尺度として採用する。高ストレスレベルの状態が慢性的に続くことで、精神障害を発症するリスクが高まるため、リカバリー経験の測定は重要である。また、開発者の Sonnentag らの実験 [13][22] においても労働生産性への影響が確認されていることから、リカバリー経験は重要な尺度と考える。

リカバリー状態 [14] は、仕事前（朝）における回復状態を評価する尺度である。仕事の後から仕事に行く前、すなわち余暇における回復の結果を表している。仕事前に、高度に回復できている状態は、身体的・心理的リソースを、仕事に多く利用可能であることを表している。逆に、回復が不十分な状態にあると、リソースが不足していることを表している。リカバリー状態は、全体的な回復、身体的な回復、精神的な回復、朝の活力について、4項目で評価する尺度である。本研究では、全体的な回復、および、ワーク・エンゲイジメントと重複する活力の2項目を省き、個別の回復状態を評価可能な、身体的な回復と精神的な回復の2項目を取り入れる。また、リカバリー状態は、朝の回復状態を評価するものであり、これには、睡眠の質も関係している [23]。そのため、後述するピッツバーグ睡眠質問票 [15] のうち、睡眠の質に関する質問を、回復状態の1項目として取り入れる。

睡眠状態は、労働生産性と関係する [14] とともに、次の日の回復状態に寄与する [23] とされている。本研究では、ピッツバーグ睡眠質問票 [15] のうち、睡眠時間に直接関係する入床時刻、入眠時刻、起床時刻に関する3項目を取り入れた。感情はポジティブ感情とネガティブ感情に分けて

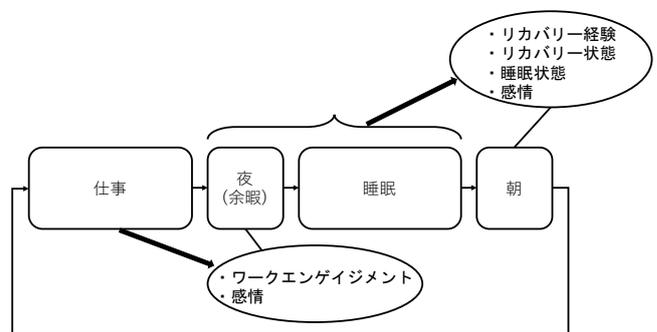


図 2 規定要因と測定タイミング

考えることができ、ポジティブ感情は、エネルギー、集中力、熱心さと正の関連を示し、ネガティブ感情は、困惑、不快と正の関連を示している [24][25]。提案手法では、スマートフォンを用いて感情を認識し、Work Attitude 評価尺度を推定することを基本アイデアとしている。また、感情尺度は、ワーク・エンゲイジメント、リカバリー経験、リカバリー状態とも関係していることが分かっている [23][26]。本研究では、UWIST[16] から 12 項目を取り入れる。

これらの尺度を用いて労働者の精神衛生をモニタリングする。図 2 に示す通り、ワーク・エンゲイジメントと仕事後の感情は、仕事が規定要因となると考えられるため、仕事後の夜（余暇）に測定を行う。リカバリー経験、リカバリー状態、睡眠状態および仕事前の感情は余暇の過ごし方と睡眠が規定要因と考えられるため、仕事前の朝に測定を行う。

3.3 スマートフォンによる収集データ

スマートフォンを用いて、活動範囲・歩数といった活動量データと、10 秒程度の労働者の自撮り動画データ、およびアンケートの収集を行う。活動量データは先行研究で頻繁に用いられており [27]、本研究においても、心理指標との相関関係があると考えられる（職場滞在時間とワーク・エンゲイジメントなど）。これらは、スマートフォンの専用ア

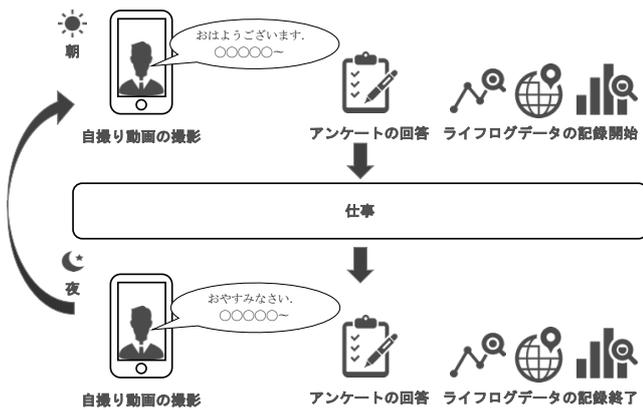


図 3 データ収集実験の流れ

アプリケーションによって自動的に収集する。

自撮り動画に関しては、先行研究 [28] において、音声情報をを用いてうつ病尺度を推定する研究が存在するが、顔画像を基に Work Attitude 関連尺度を評価する研究は確認できていない。また、近年、顔画像から表情を認識し、感情を推定する研究 [29] が行われており、自撮り動画を処理することで、撮影時における顔画像による感情認識も行える。

さらに、本研究では、自撮り動画の撮影時に、今の気持ちを一言メッセージとして発言してもらおう。これにより、発話内容からも感情推定および特徴量の抽出を行う。このような音声と顔画像を含んだマルチモーダルな情報である動画データを収集し、音声から得られる特徴量、顔画像から得られる特徴量、発話内容から得られる特徴量を組み合わせることで日常的な心理尺度を推定する手法は、これまで行われておらず、音声だけでは明らかにならなかった心理尺度との相関関係が明らかになると考えられる。

4. データ収集実験

4.1 実験方法

図 3 に実験の流れを示す。被験者は、朝起きてから仕事へ向かうまでに自撮り動画の撮影とアンケートの回答を行い、ライフログデータの記録を開始する。仕事から帰宅した後、就寝前に自撮り動画の撮影とアンケートの回答を行い、ライフログデータの記録を終了する。撮影に際しては、「おはようございます」と「おやすみなさい」を朝夜それぞれ収録し、ベースタームとする。これにより同じタームに対する声質の違いを比較可能とする。さらに、ベースタームの後に一言メッセージ (ex. 「今日はよく寝た。元気出していこう」「まだ眠い。会社行きたくないな〜」「今日はお客さんに褒められた。よく頑張った。」「課長に怒られた。寝て忘れよう」など) を発言してもらおう。

4.2 実験結果

表 2 に得られたデータの概要をまとめる。9 人の被験者に対し、2 週間のうち休日を除いた 10 日分のデータを収集

表 2 データ概要

実験期間	2 週間
人数	9 人 (男のみ)
収集データ	動画 アンケート GPS 活動量
データ数	朝:78 セット 夜:77 セット

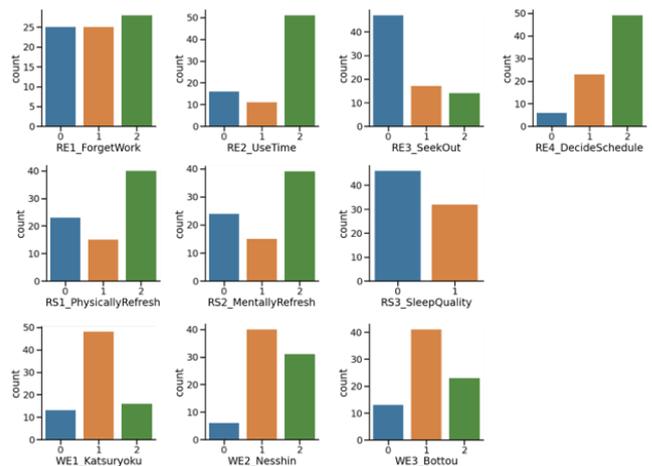


図 4 データ加工後のアンケート回答結果

した。アンケートの回答忘れや動画の撮り忘れなどの影響により、推定モデルの構築に使用するデータセットは朝 78 セット夜 77 セットの合計 155 セットとなった。

目的変数となるアンケート回答結果の分布は、全体的な傾向として 0 と回答したクラスの度数が小さく、RE4 に関しては 1 であった。目的変数となるアンケート回答結果が不均衡かつ、著しくデータ数が少ないクラスが存在し、評価検証が難しいため、5 段階、7 段階の評価項目は 3 段階に、4 段階の評価項目は 2 段階に変更して検証を行う。データ加工後のアンケート回答結果を図 4 に示す。

5. Work Attitude 推定モデルの構築と評価

5.1 特徴量抽出

特徴量の一覧を表 3 にまとめる。動画データからは画像、音声、テキスト (発話内容) の 3 つのチャンネルに切り分け、それぞれのチャンネルから得られる感情値を特徴量とした。感情値の取得にはそれぞれ Face API (Microsoft 社)、Empath API (Empath 社)、Cotoha API (NTT コミュニケーションズ社) を使用した。なお、画像チャンネルの特徴量に関しては、動画から 5 枚の画像を抽出してそれぞれの感情値の平均値、中央値、分散、最大値を特徴量としている。

GPS データからは 1 日の総移動距離を、活動量データからは 1 日の歩数の合計をそれぞれ算出し特徴量として用いる。アンケートについては加工後のデータを目的変数とし

表 3 特徴量一覧

チャンネル名	特徴量	備考
画像チャンネル	smile	
	anger	
	contempt	
	disgust	0 から 1 までの数値.
	fear	1 に近いほどその感情である
	happiness	可能性が高い
	neutral	
	sadness	
音声チャンネル	surprise	
	calm	
	anger	0 から 50 までの数値.
	joy	50 に近いほどその感情である
	sorrow	可能性が高い
テキストチャンネル	energy	
	positive/neutral/negative	
	P	positive/neutral/negative は
	N	テキスト全体としての気分判定
	PN	を行う. 数値化にあたり
	悲しい	それぞれ 0,1,2 として対応付けた.
	不安	その他はテキスト内に該当する
	好ましい	感情の言葉があった場合カウント
	興奮	する
	喜ぶ	例) 今日は楽しい 1 日だった.
安心	→判定:positive (数値では 0)	
嫌	好ましい=1	
GPS	Distance	総移動距離 [m]
活動量	Step	歩数 [歩]
アンケート	Sleep Efficiency	睡眠効率 [%]

表 4 正解値の分散および特徴選択前後の F 値と特徴量選択した際の特徴量数

項目	分散	Random Forest			XGBoost		
		全特徴量	選択後	選択個数	全特徴量	選択後	選択個数
RE1	0.83	0.35	0.59	5	0.41	0.47	4
RE2	0.82	0.39	0.58	1	0.34	0.63	7
RE3	0.78	0.27	0.33	8	0.27	0.35	14
RE4	0.64	0.41	0.45	13	0.36	0.44	4
RS1	0.88	0.40	0.47	2	0.45	0.43	10
RS2	0.88	0.39	0.52	1	0.42	0.59	8
RS3	0.5	0.43	0.59	6	0.46	0.58	8
WE1	0.62	0.34	0.38	8	0.36	0.46	1
WE2	0.62	0.46	0.47	7	0.35	0.43	11
WE3	0.68	0.32	0.41	4	0.42	0.51	7

て用いるが、睡眠に関する項目のみ、睡眠効率を算出して特徴量として用いた。

5.2 学習モデル構築

ワークエンゲイジメント、リカバリー経験、リカバリー状態の各回答結果を目的変数として、ランダムフォレスト、XGBoost を用いて学習モデルの構築を行った。表 4 に項目ごとの正解値の分散、特徴量選択前後の F 値、特徴量選

択時の特徴量の数を示す。また、各項目ごとの重要度の高い特徴量をまとめたものを表 5 に示す。

5.3 考察

表 4 より、学習に用いる特徴量を選択したことで精度が向上することを確認した。特徴選択後に最も精度が向上したのは Random Forest アルゴリズムでは RE1 で 24%、XGBoost アルゴリズムでは RE2 で 29%精度が向上した。

表 5 各項目における特徴量の重要度トップ 5

	Random Forest				
	Top1	Top2	Top3	Top4	Top5
RE1	Sleep_efficiency	neutral_m	joy	calm	Distance
RE2	sadness_v	sadness_max	sadness_m	Sleep_efficiency	sadness_med
RE3	Distance_log	Distance	Sleep_efficiency	neutral_v	sadness_v
RE4	Sleep_efficiency	neutral_max	neutral_med	sadness_m	sadness_med
RS1	neutral_med	neutral_max	sadness_m	Sleep_efficiency	neutral_m
RS2	sadness_m	calm	sadness_max	joy	neutral_max
RS3	Sleep_efficiency	Step	Distance	Step_log	Distance_log
WE1	Distance	Distance_log	sadness_max	sadness_m	sadness_v
WE2	happiness_v	smile_v	happiness_m	smile_max	happiness_max
WE3	calm	sadness_v	Step_log	sadness_max	Step
	XGBoost				
	Top1	Top2	Top3	Top4	Top5
RE1	Sleep_efficiency	smile_v	smile_max	smile_med	sadness_v
RE2	sadness_max	sadness_med	N	neutral_m	sadness_v
RE3	disgust_m	sadness_m	fear_med	fear_m	surprise_med
RE4	neutral_max	joy	surprise_max	neutral_m	surprise_m
RS1	neutral_med	disgust_v	fear_max	smile_max	neutral_max
RS2	N	fear_med	fear_v	P	neutral_m
RS3	fear_med	surprise_med	anger_v	Distance	P
WE1	sadness_max	surprise_v	neutral_v	fear_v	sadness_m
WE2	smile_v	energy	P	surprise_m	smile_m
WE3	sadness_max	kanashii	surprise_v	disgust_m	energy

Random Forest アルゴリズムでは WE2 に、XGBoost アルゴリズムでは RS1 に精度の向上が認められなかった。全体では、機械学習アルゴリズムの違いによらず、精度は平均で特徴量選択前後で 10% 向上した。ドメインごとに見ると、改善率が最も高かったのは Random Forest アルゴリズムでは RE で 13%、XGBoost アルゴリズムでも RE で 13% の改善であった。また、どちらの機械学習アルゴリズムを用いた場合でも、RE3 の精度が低く、特徴量選択後、かろうじて 30% に届くことが確認された。

また、特徴量の重要度は、全体的には Random Forest アルゴリズムでは、睡眠効率や一日の総移動距離といった、ライフログモーダルの特徴量の重要度が高い傾向にあることが確認できた。XGBoost アルゴリズムでは、対照的にライフログデータに関する特徴量は一つも入っておらず、悲しみや驚きの表情といった、画像モーダルの特徴量の重要度が高い傾向にあることが確認できた。

各項目については、2つのアルゴリズムを比較して、精度が高い方に着目して考察する。RE1 は睡眠効率の重要度がもっとも高く、RE2 は悲しみの表情の重要度、RE3 は困惑の表情、RE4 は睡眠効率の重要度がもっとも高かった。リカバリー経験のドメイン全体としては、重要度がもっとも高い特徴量は、ライフログモーダルと画像モーダルの特徴量が同数となった。RS1 はニュートラルな表情の重要度が最も高く、RS2 はネガティブワードの重要度が最も高かった。RS3 は睡眠効率の重要度が最も高かった。リカバリー状態のドメイン全体としては、画像、テキスト、ライフログの 3つのモーダルの特徴量が同数となった。WE1 は悲

しい表情の重要度が最も高く、WE2 は幸せな表情の重要度が最も高く、WE3 は悲しい表情の重要度が最も高かった。ワーク・エンゲイジメントのドメイン全体としては、画像モーダルの特徴量のみが最も重要な特徴量であった。この結果より、Work Attitude の測定のためにはマルチモーダルな情報が必要であることが考えられる。

6. おわりに

本研究では、労働者が働き方を日常的かつ客観的に振り返るために、スマートフォンから得られるデータを用いて Work Attitude を推定し、労働者にフィードバックする「Work Attitude PLR 収集基盤」の構築を提案した。提案手法を実現するためには、労働者の仕事に対する状態を日常的かつ包括的に測定する Work Attitude 評価尺度を開発し、スマートフォンから得られるデータによってその評価値を推定するモデルを構築する必要がある。そして、推定モデルの構築のためには、推定精度や必要となる説明変数を明らかにする必要がある。本手法の実現可能性を評価するため、Work Attitude 評価尺度のパイロットスケールを開発し、その尺度を用いてデータ収集実験を行った。データ収集実験では、10 日間に渡り 9 人の被験者から動画データ及びライフログデータとアンケートを収集した。その後、収集したデータからデータセットを作成し、評価実験を行った。

実験の結果より、特徴量選択を行うことで平均 51% の F 値で Work Attitude の各項目を推定できることがわかった。また、Work Attitude の推定には、有効な特徴量のモー

ダルが異なることがわかった。

今後の展望として、推定モデルの精度が 50%と低いのはデータ数の不足が原因として考えられるため、大規模なデータ収集実験を実施することが挙げられる。

謝辞 本研究は、JSPS 科研費 JP19K11924 の助成を受けるとともに、大阪大学グランドチャレンジ研究により大阪大学ライフデザイン・イノベーション研究拠点から委託されたものです。

参考文献

- [1] 厚生労働省. 平成 30 年「労働安全衛生調査 (実態調査)」の概況. 2018.
- [2] T. Yamauchi, T. Yoshikawa, M. Takamoto, and et al. Overwork-related disorders in japan: recent trends and development of a national policy to promote preventive measures. *Industrial Health*, Vol. 55, pp. 293–302, 2017.
- [3] 高橋正也. 過労死等の実態解明と防止対策に関する 総合的な労働安全衛生研究. p. 23=35, 2017.
- [4] 日本生産性本部. 労働生産性の国際比較 2018. 2018.
- [5] 厚生労働省. 労働安全衛生法に基づくストレスチェック制度実施マニュアル. 2019.
- [6] V.P. Cornet and R.J. Holden. Systematic review of smartphone- based passive sensing for health and well-being. *Journal of Biomedical Informatics*, Vol. 77, pp. 120–132, 2018.
- [7] C. Amenomori, T. Mizumoto, and et al. A method for simplified hrqol measurement by smart devices. *Wireless Mobile Communication and Healthcare*, Vol. 247, pp. 91–98, 2018.
- [8] Jaques Natasha and et al. Predicting students’ happiness from physiology, phone, mobility, and behavioral data. *Int Conf Affect Comput Intell Interact Workshops*, 2015.
- [9] A.J. Sano, A.Z. Phillips, and et al. Recognizing academic performance, sleep quality, stress level, and mental health using personality traits, wearable sensors and mobile phones. *Proc. IEEE 12th International Conference on Wearable and Implantable Body Sensor Networks (BSN)*, pp. 1–6, 2015.
- [10] M. Boukhechba, A. Daros, and et al. Demonic salmon: Monitoring mental health and social interactions of college students using smartphones. *Smart Health, CHASE 2018 Special Issue*, Vol. 9-10, pp. 192–203, 2018.
- [11] Y. Fukazawa, T. Itoh, and et al. Predicting anxiety state using smartphone-based passive sensing. *Journal of Biomedical Informatics*, Vol. 93, pp. 103–151, 2019.
- [12] W. B. Schaufeli, A. Shimazu, J. Hakanen, M. Salanova, and H. De Witte. An ultra-short measure for work engagement: The uwe-3 validation across five countries. *European Journal of Psychological Assessment*, Vol. 35, pp. 577–591, 2019.
- [13] S. Sonnentag and C. Fritz. The recovery experience questionnaire: Development and validation of a measure for assessing recuperation and unwinding from work. *Journal of Occupational Health Psychology*, Vol. 12, No. 3, pp. 204–221, 2007.
- [14] C. Binnewies, S. Sonnentag, and E. J. Mojza. Daily performance at work: feeling recovered in the morning as a predictor of day-level job performance. *Journal of Organizational Behavior, J. Organiz. Behav.*, Vol. 30, pp. 67–93, 2009.
- [15] D. J. Buysse, C. F. Reynolds, T. H. Monk, S. R. Berman, and D. J. Kupfer. Pittsburgh sleep quality index (psqi). *Psychiatry Research*, Vol. 28, pp. 193–213, 1989.
- [16] M. Gerald, M. J. Dylan, and A. C. Graham. Refining the measurement of mood: The uwest mood adjective checklist. *British Journal of Psychology*, Vol. 81, pp. 17–42, 1990.
- [17] 島津明人. 職業性ストレスとワーク・エンゲイジメント. *ストレス科学研究*, Vol. 25, pp. 1–6, 2010.
- [18] W.B. Schaufeli, M. Salanova, V Gonzalez-Romá, and et al. The measurement of engagement and burnout: A two sample confirmative analytic approach. *Journal of Happiness Studies*, Vol. 3, pp. 71–92, 2002.
- [19] W. B. Schaufeli and A.B. Bakker. Job demands job resources and their relationship with burnout and engagement :a multi-sample study. *Journal of Organizational Behavior*, Vol. 25, pp. 293–315, 2004.
- [20] W. B. Schaufeli and A.B. Bakker. The conceptualization and measurement of work engagement:a review. *Work Engagement: A Handbook of Essential Theory and Research*, pp. 10–24, 2010.
- [21] A. Shimazu, W.B. Schaufeli, S. Kosugi, and et al. Work engagement in japan:validation of the japanese version of utrecht work engagement scale. *Applied Psychology: An International Review*, Vol. 57, pp. 510–523, 2008.
- [22] A. Shimazu, S. Sonnentag, K. Kubota, and N. Kawakami. Validation of the japanese version of the recovery experience questionnaire. *Journal of Occupational Health*, Vol. 54, pp. 196–205, 2012.
- [23] S. Sonnentag, C. Binnewies, and E. J. Mojza. “did you have a nice evening?” a day-level study on recovery experiences, sleep, and affect. *Journal of Applied Psychology*, Vol. 93, No. 3, pp. 674–684, 2008.
- [24] D. Watson, L.A. Clark, and A. Tellegen. Development and validation of brief measures of positive and negative affect: The panas scales. *Journal of Personality and Social Psychology*, Vol. 54, pp. 1063–1070, 1988.
- [25] 川人潤子, 大塚泰正, 甲斐田幸, 中田光紀. 日本語版 the positive and negative affect schedule (panas) 20 項目の信頼性と妥当性の検討. *広島大学心理学研究*, Vol. 11, pp. 225–240, 2011.
- [26] S. Sonnentag, K. Eck, C. Fritz, and J Kühnel. Morning reattachment to work and work engagement during the day: A look at day-level mediators. *Journal of Management*, Vol. 46, No. 8, pp. 1408–1435, 2020.
- [27] Y. Fukazawa, N. Yamamoto, T. Hamatani, and et al. Smartphone-based mental state estimation: A survey from a machine learning perspective. *Journal of Information Processing*, Vol. 28, pp. 16–30, 2020.
- [28] A.A. Farhan, J. Lu, J. Bi, A. Russell, B. Wang, and A. Bamis. Multi-view bi-clustering to identify smartphone sensing features indicative of depression. *IEEE 1st International Conference on Connected Health: Applications, Systems and Engineering Technologies (CHASE)*, pp. 264–273, 2016.
- [29] F. Noroozi, M. Marjanovic, A Njegus, and et al. Audio-visual emotion recognition in video clips. *IEEE TRANSACTIONS ON AFFECTIVE COMPUTING*, Vol. 10, No. 1.