

1 対多遠隔指示の指示者支援技術－作業者の戸惑い・順調レベル提示による負荷の低減

代表研究者	王 子 洋	産業技術総合研究所 人間情報インタラクション研究部門 研究員（前職：北陸先端科学技術大学院大学 助教）
共同研究者	李 秀 永	北陸先端科学技術大学院大学 先端科学技術研究科 博士 後期課程学生
共同研究者	大 槻 麻 衣	産業技術総合研究所 人間拡張研究センター 主任研究員
共同研究者	葛 岡 英 明	東京大学大学院情報理工学系研究科 教授
共同研究者	由井 蘭 隆 也	北陸先端科学技術大学院大学 先端科学技術研究科 教授

1 はじめに

近年、製造業をはじめ、産業はインテリジェントな機械やロボットに大量に使用し始めているが、作業者の役割は依然として重要である。作業者の作業を支援するために、専門家や経験豊富な指示者が遠隔で作業者に指示を与える遠隔コラボレーションの重要性は十分に示唆されている。しかし、現場の労働力不足のため、遠隔コラボレーションをより効率的に行う必要性が高まっている。例えば、1人の専門家が複数の新人作業者を並列に指示する1対多の遠隔コラボレーションのニーズが増えてくる。一対一の遠隔コラボレーションと異なり、1対多の遠隔コラボレーションでは、指示者は複数の作業者をサポートする際に非常に高い作業負荷を感じていることが分かっている[1]。そこで、一対多の遠隔コラボレーションを円滑に実現するために、本研究は指示者の作業負荷を軽減することを目的とした。

作業負荷が高い主な原因の1つは、指示者が効果的に複数の作業者を支援するために、作業員の状態を同時に監視しなければならないことである。このようなマルチタスクでは、人の精神的負荷が高いことが知られている。例として、Fanらは、参加者に対して、複数のメーターの異常検知の課題を課し、監視対象を増やせば増やすほど、主観的な負荷は有意に高くなった。一方、反応時間も増加し、異常値の発見精度は大幅に低下したことを示した[2]。そのため、指示者の負荷を下げる重要なアプローチの1つは、作業者を常に監視しなくても、指示者が作業者の状態を把握できるよう支援することである。

人間工学では、人間の状態を推定する研究が長い歴史を持っている。そのうち、生理状態と心理状態に分けられる。前者では、作業者の身体を分析し、正しい姿勢や転倒など危険な動きを行っているか検出する研究が多い（例：[3, 4]）。例えば、Nwakanmaらは、LIDARや振動センサーなどのセンサーを用い工場内の緊急な状態を検出した[5]。一方、後者では、作業者の心理状態の把握においても重要な役割を果たす。例えば、作業者の精神的な負荷を評価することで、業務の負荷・スケジュールの管理につながる。過去の関連研究では、疲労[6, 7]、感情[8]、注意[9]、認知的負荷[10, 11, 12]など、心理状態のさまざまな側面が注目されていた。

遠隔コラボレーションでは、指示者は作業者が支援を求める際に、円滑に対応するために、その場の作業者の作業進捗と問題点を把握する必要がある。本来指示者は作業者が支援を求めるタイミング（下記、支援タイミングとその直前の情報を把握すれば円滑にできるが、支援タイミングがわからないので、常に作業者に監視をしなければならないことになる。この観点から、支援タイミングとその前の一定間隔の作業を指示者に提示すると、指示者が作業員の監視から解放できると考えられる。

こういう機能を実現するために、支援タイミングを検出する必要がある。本研究では、支援タイミングに関わる二つの心理状態、すなわち「戸惑い度」と「順調度」に焦点を当てた。戸惑いとは、矛盾、葛藤、誤った情報によって引き起こされる認知的不調和の状態である[13, 14]。組み立て作業では、指示を理解するのが難しい場合や、組み立てるのが難しい場合など、さまざまな状況で戸惑うことがある。戸惑い度が高ければ高いほど、指示者に支援を求めたくなると考えられる。順調度とは、タスクが自分の期待よりも早く完了しているかどうかを表す感情である。順調度が高ければ、満足感や自信につながり、作業に対してポジティブな影響をもたらす一方、指示者に支援を求めたい考えが少ないと考えられる。

本報告書では、1対多遠隔コラボレーションの指示者の負荷を下げるシステムを開発するため、作業者の戸惑い程度と知覚した作業のスムーズさの検出手法を提案する。検出手法を開発するにあたり、組み立て作

業の実験を実施し、戸惑い程度と知覚した作業のスムーズさの自己評価アンケートの結果と、作業者の筋電図（EMG）変化、視線情報の関連性を報告する。

2 戸惑い程度と知覚した作業のスムーズさの検出手法の開発用実験設計

2-1 タスク

遠隔コラボレーションに関わる作業は組み立て作業、修理作業、メンテナンスタスクなどが含まれる。本実験でのタスクは、遠隔協調作業研究で特によくみられる様々なパーツを組み立てる作業を採用した。組み立て作業を模擬するために、2つの自動車組み立てキット（それぞれ150個の部品を含む）を使用し、3つの組み立てタスクを作成した。各タスクには約25ステップが含まれ、ステップごとに1ページのマニュアルを作成した。マニュアルには、部品の種類と数、組み立て対象の図が含まれていた（図1）。組み立て時に戸惑いを生じさせよう、異なる部品や一致しない対象物の図などの情報を故意に入れた。

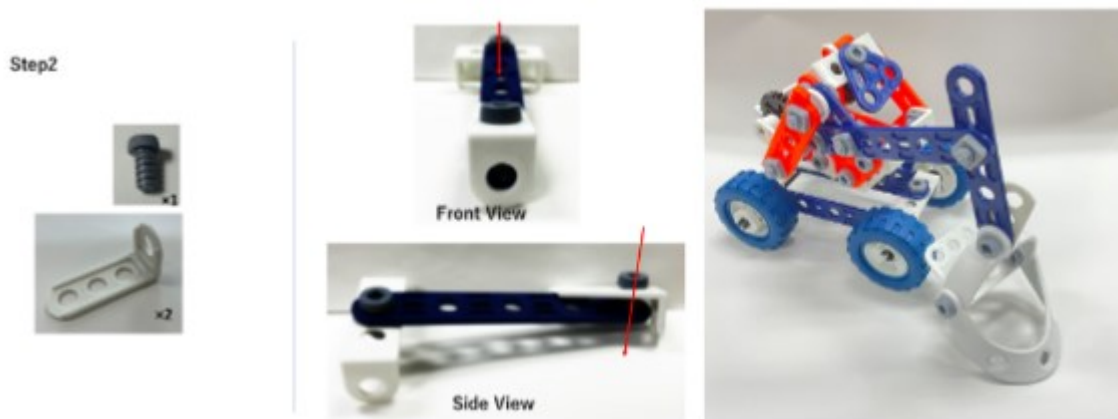


図1：実験に使用したマニュアルのサンプル

環境については、組み立て作業は120 cm x 80 cmの作業台上で行った。作業台はテーブルで3つのエリアに分けられ、未使用の部品は右のエリアに置かれ、マニュアルは左のエリアに置かれ、真ん中の部分は組み立て作業用エリアであった。

2-2 実験参加者

この実験に参加する40人の実験参加者は、北陸先端科学技術大学院大学の学生から募集された。実験参加者全員から北陸先端科学技術大学院大学が承認したインフォームドコンセントを取得された。平均年齢は27.7歳（SD = 3.3）で、うち25人が男性であった。参加者は全員中国語を母国語とし、謝金として一人あたり3000円を受け取った。

2-3 実験手順

まず、実験参加者は実験室に入り、実験の説明を聞いた後、インフォームドコンセントに署名した。その後、眉間筋に筋電図表面電極を装着し、ARのヘッドマウントディスプレイ Magic Leap 2 を装着した。EMGと Magic Leap 2 のキャリブレーション後、実験参加者が手順を理解したかを確認するために、組み立て作業の練習セッションが開始された。その後、組み立て作業セッションを2回行った。作業中、実験参加者はステップずつ、マニュアルのページをめくり、部品エリアから部品を探し、組み立てて組み立て作業を行った。ステップを終えた後、実験参加者は、Magic Leap 2 に表示された戸惑い度と順調度の質問をキーボードで回答した。

2つの質問に答えた後、参加者は次のページをめくり、次のステップを進めた。各セッションは、参加者がすべてのステップを完了するか、50分の制限時間に達した時点で終了した。参加者は2つのセッションの間に15分間の休憩が含まれた。

2-4 計測項目

(1) 主観評価

各ステップで経験した戸惑い度と順調度を測定するために以下の質問を回答させた。前者の質問は5段階評価、後者の質問は7段階評価でした。

- 質問1：このステップを行う際に、どの程度の戸惑いを感じたか？
- 質問2：このステップを完了するために費やした時間は、あなたの予想と比べて早かった？

(2) 客観的尺度

作業者の戸惑い度と順調度を客観的に推定するために、作業者の視線と筋電を記録した。視線では、Magic Leap2のカメラと頭部座標と視線記録機能を用い、作業者の各ステップの作業状態の時系列変化を検出した。具体的にマニュアル理解状態、組み立て状態、および部品探し状態であった。そして、作業状態の時系列変化に基づき、状態変化数や各状態の数など指標を集計した。筋電について、関連研究で皺眉筋と大頬骨筋から戸惑いを推定でき、そのうち皺眉筋の精度がより高いと示されている[16]。そこで、皺眉筋の筋電について、Shimmer社のセンサーキットを用い記録した。また、各ステップの経過時間も録画データから算出した。

3 結果

3-1 主観評価と経過時間の関係

各ステップの2つの主観評価の回答と経過時間との相関を算出した(表1)。参加者は各ステップのアンケートに回答しているため、参加者内効果の影響を考慮し、反復測定相関を適用した[15]。その結果、3つの変数間に中程度の相関があることが示された。

表1 各ステップの2つの主観評価の回答と経過時間との相関

	戸惑い度	順調度	時間
戸惑い度	1	-0.52	0.47
順調度	-0.52	1	-0.55
時間	0.47	-0.55	1

3-2 皺眉筋筋電(EMG)

3-2-1 皺眉筋筋電の相関

Python バイオシグナル処理ライブラリである NeuroKit を用い、EMG データを分析した。まず、各ステップの512HzのEMG生信号を、ライブラリ biosppy の前処理法を用いてEMG信号の振幅を求めた。次に、同ライブラリの手法でEMG信号の振幅からEMGイベントを検出した。この手法では、EMG信号の振幅が平滑化および全波整流されたEMG信号の絶対値の平均の1.2倍を超える場合にEMGイベントと判断された。最後に、EMGイベントの数を集計した。

図2は、EMGイベントの数の分布を示しており、各色は異なる知覚レベルの混乱を表している。繰り返し測定と相関分析により、戸惑い度と中程度の相関があることが示された($r=0.32$, $p<.001$)。さらに、混合モデル線形回帰が構築された。EMGイベントの数を説明変数とし、戸惑い度を予測変数とし、実験参加者とセッションをランダム因子とした。その結果、知覚された混乱のレベルに対するEMGイベントの数の有意な影響が示された($\beta=0.004$, $SE=.0004$, $p<.0001$)。

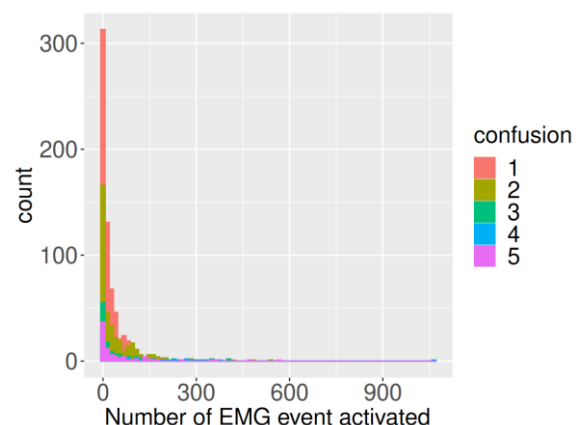


図2. 各ステップのEMGイベント数のヒストグラム

3-2-2 皺眉筋筋電（EMG）による機械学習モデル

皺眉筋筋電に基づき戸惑い度を推定するモデルを開発するために、EMG 周波数データをスペクトログラム（Spectrogram）画像へ変換し、CNN モデルによる予測と分類を試みた。スペクトログラムは、時間とともに変化する音や波を視覚化してグラフで表現する手法である。一般的に X 軸（時間）と Y 軸（周波数帯）の 2 次元ピクセルデータで表現される。

スペクトログラムのための信号処理には代表的に高速フーリエ変換（fast Fourier transform; FFT）が使用されるが、信号全体の値を正弦波の和で表す FFT は、信号全体の中の急激な変化がどこで起こったのかを効率的に表現することが難しい。そのため、分析に用いたスペクトログラム画像は、連続ウェーブレット変換（continuous wavelet transform; CWT）によって処理された。CWT は、任意の信号をウェーブレットという一定の形を持つ小さい波形で分解することで、周波数の特徴を時間に沿って分析できるようにする長所を持つ手法である。

データの処理と分析に使用されたコンピューティング環境は、Apple Mac Studio (Apple M2 Max, 32GB Ram), macOS Ventura 13.6.7 であり、使用されたソフトウェアとライブラリは Python 3.11.6 と PyTorch 2.2.2, NeuroKit2 0.2.9, Polars 0.20.31, Ssqueezepy 0.6.5 である。また、PyTorch で作られたモデルには Apple の Metal Performance Shaders (MPS) による GPU アクセラレーションテンソル計算を適用した。また、データの前処理のために、スペクトログラム画像を実験参加者の情報と戸惑い度の回答に紐付けて生成する並列分散処理関数を実装した。

● 教師あり手法によるアプローチ

二次元 EMG スペクトログラム画像と実験参加者の戸惑い度の回答に基づいた特徴抽出と分類は、画像処理に特化された CNN による手法を採択した。まず、各ステップの EMG データを前述の前処理で EMG 信号の振幅を求め、その後、そのデータを用いて Ssqueezepy ライブラリによる CWT 処理を行い、最終的に生成されたスペクトログラムデータを保存した。生成されたスペクトログラム画像の例は図 3 である。

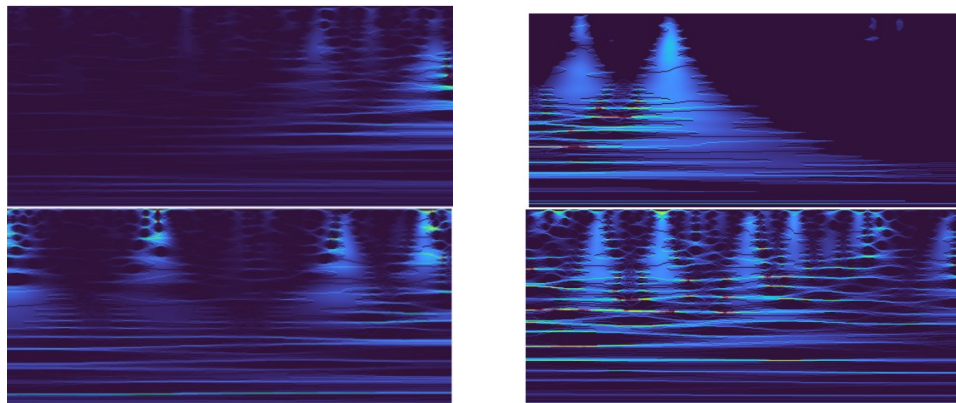


図 3. 左図：実験参加者 3 の戸惑い度回答 1（上）、戸惑い度回答 4（下）の EMG スペクトログラム／
右図：実験参加者 10 の戸惑い度回答 1（上）、戸惑い度回答 4（下）の EMG スペクトログラム

分類に使われた CNN モデルは、Pytorch ライブラリの PyTorch によるニューラルネットワークのベースクラス（torch.nn.Module）を用いて実装された。モデルは 224 x 224 px の画像を入力とする 3 つの 2D データ畳み込み層（Conv2d）と活性化関数 ReLU、3 つの最大値プーリング層、3 つの全結合層で構成された。ネットワークの全体構造は図 4 に示した。

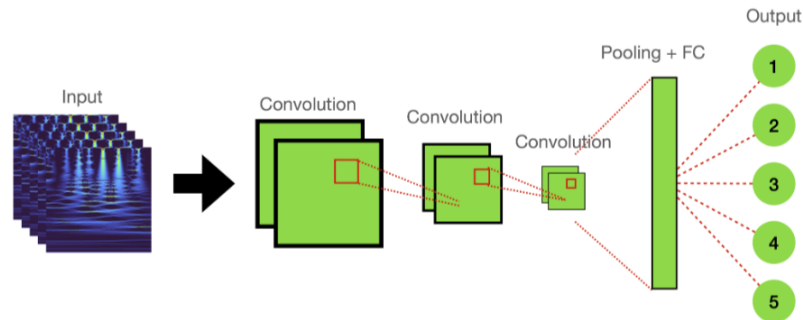


図 4. 戸惑い度推定に使われた CNN モデル構造

CNN による戸惑い推定の可能性を確認するために、実験参加者内の分析を行った。分析では、各々の実験参加者ごとのタスクデータを、3割をテストデータとして分離し、残りの7割のデータをモデルの学習に用いて、自分のデータだけで学習されたモデルによる検証を行った。各々の実験参加者ごとのテストデータに対する推定の平均正解率は 41.73% であり、全体実験参加者の学習データに対するテストデータの推定の平均正解率は 41.54% であった。

その結果に対して、タスク全体のスペクトログラム画像を用いて5段階の戸惑い程度を推定することは難しいだと判断し、戸惑い有り無しとの二値推定を試した。また、全体のデータを2つのグループに分割した k 分割交差検証による学習と検証を行った。5段階の戸惑い程度推定では全体 41.50% の平均正解率が、戸惑いの有り無しとの二値推定では全体 54.20% の平均正解率が確認された。

● 教師なし手法によるアプローチ

実験で収集された実験参加者の戸惑い程度に対する回答は、その程度や有無に個人差を含む主観的な影響が存在する可能性がある。したがって、EMG 信号に基づく戸惑い度の推定方法の一つとして、信号を一定間隔の窓関数で分割し、それを2次元スペクトログラム化したデータを教師なしCNNで分類する手法を試みた。具体的には、CNN で入力値の最終分類のために使用される Fully Connected Layer に入力される特徴ベクトルだけを抽出し、それを K-means クラスタリングを用いて類似度を求め、分類する方法を適用した。

K-means クラスタリングは、K 個の任意で指定されたクラスター中心 (Centroid) に基づいて、分類したいデータを最も近いクラスター中心に配置し、クラスター中心を再設定する処理を繰り返し、データを分類するアルゴリズムである。実際にデータを分類する前に、最適なクラスター数を求めるために、すべてのデータを用いて、各 K 値に対する Inertia 値を求める。Inertia 値は、各クラスターの中心からクラスターに配置されたデータ間の距離を合計した値で、この値が小さいほどクラスターの凝集度が高いと言える。

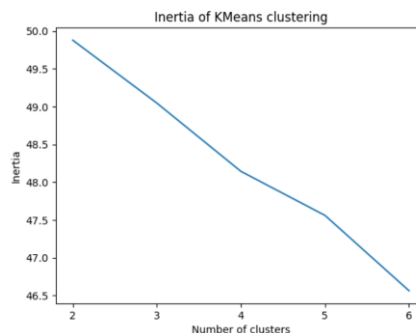


図 5. K-means Inertia 値

各 K 値に対する Inertia 値を 2 から 6 まで計算した結果、図 5 に示したように、K が 4 の場合、凝集度減少の効果が少なくなることが確認でき、4 をクラスター数として扱った。

特徴抽出には、CNN モデルを用いて分類を行った。前述のように、学習データを CNN に投入し、最終層のベクトルデータを K 値が 4 である K-means クラスタリングに投入した。抽出された特徴量は、定性的な評価を通じて、最も範囲の小さい特徴 1 を分類対象とした。

表 2 特徴選択

特徴 No.	平均	最小	最大	範囲
0	0.746	0.717	0.781	0.064
1	0.758	0.737	0.777	0.04
2	0.744	0.718	0.775	0.057
3	0.753	0.715	0.79	0.075

上記の通り実装された分類機を用い、実験参加者間の分析を行った。全体実験参加者のデータを均等にグループ 1 とグループ 2 に分け、各々のグループを k 分割交差検証による分類機の学習と検証をした。分類機の正解を判別する基準は、回答された戸惑い度が 2 以上の場合は 4 個以上の戸惑った区間が検出された時を True Positive とし、回答された戸惑い程度が 1 の場合に 4 個未満の戸惑った区間が検出された時を True Negative として扱った。その結果、グループ 1 のデータで学習された分類モデルの f1 スコアは 60.74 であったが、グループ 2 のデータで学習された分類モデルの f1 スコアは 1.78 しかなかった。

3-3 視線情報による作業状態

作業状態に関する指標と

前述のように視線と頭部座標に基づき作業者の時系列の状態変化を検出した。そして、状態の数、各状態の数、経過時間などの 29 種の指標を算出した。これらの指標を特徴として戸惑い度と順調度を推定する SVM モデルを構築した。戸惑い度の推定モデルでは、正解率が 64.4%であり、順調度の推定モデルでは、正解率が 43.2%であった。各クラスは学習データにおいて不均衡であるため、加重平均正解率が算出された。その結果、戸惑い度の推定モデルの加重平均正解率が 29.1%であった。また、順調度の推定モデルの加重平均正解率が 25.4%であった(図 6)。

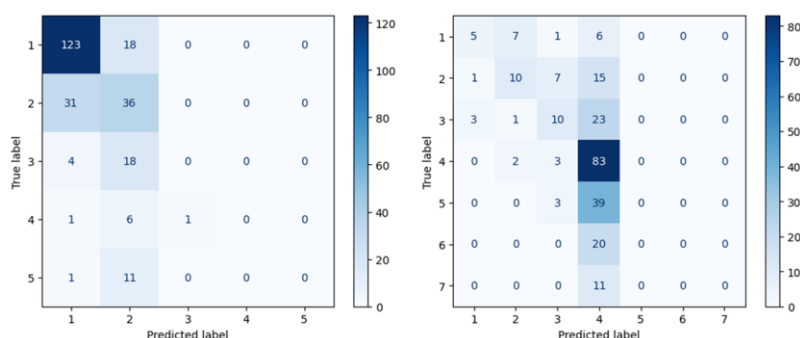


図 6 左：SVM による 5 クラスの戸惑い度の混同行列／右：SVM による 7 クラスの順調度の混同行列

戸惑い度が 2 以上を戸惑い有、1 を戸惑い無とする二値分類モデルでは、正解率は 79.2%，加重平均正解率は 79%であった。順調度が 4 以上を順調、3 以下を順調しないとする二値分類モデルでは、正解率は 79.2%，加重平均正解率は 75.3%であった(図 7)。

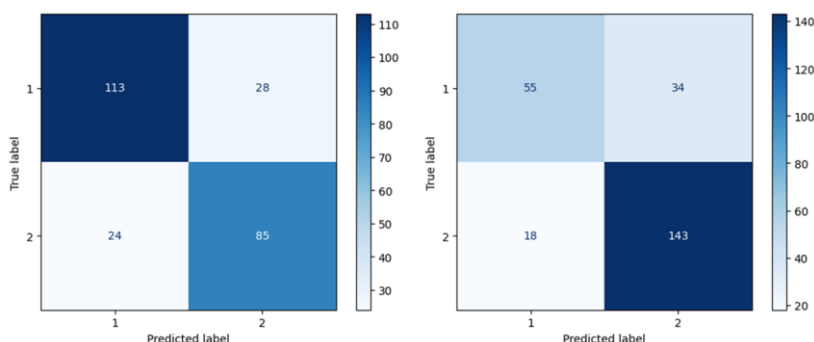


図 7 左：SVM による 2 クラスの戸惑い度の混同行列／右：SVM による 2 クラスの順調度の混同行列

4 考察と結論

本研究では、1 対多の遠隔コラボレーションの効率を高めるために、指示者の作業者を監視する負荷を下げる手法の開発を目的とした。これを実現するために、作業者の戸惑い度と順調度を検出できるモデルを開発した。1 対多の遠隔コラボレーションでよく見られる組み立てタスクの模擬実験を行い、作業者の視線と皺眉筋筋電 (EMG) のデータを収集した。

皺眉筋筋電について、生の EMG 信号から検出された筋電のイベントの数が作業者の戸惑い度と相関がみられた。筋電信号の周波数と時間で生成された二次元のスペクトログラム画像を特徴とし、戸惑い度を推定する教師あり CNN モデルを開発した。5 段階の戸惑い度を推定する場合、正解率が 41.54%であった。一方、戸惑いの有無の二値分類では 54.20% の平均精度が確認された。一方、主観評価の戸惑い度を用いない教師なし手法では、二次元筋電スペクトログラム画像を CNN 手法で特徴抽出し、K-means クラスタリング手法を試みた結果、学習データにより精度が大きく変わったことを示した。これらの試みにより、皺眉筋筋電のイベントが戸惑い度と相関があるが、推定モデルの特徴として用いるのは限界がある。一つの原因は、筋電センサーの干渉問題である。本研究では、実験参加者の眉の周辺に電極を付けた一方、AR のヘッドマウントディスプレイ Magic Leap2 も同じ場所に接触した。この場合、Magic Leap2 で固定されたので、電流の計測に影響を与えうる。さらに、実験参加者が自然に表情を表せない可能性が高い。今後の課題として、精度が高いアルゴリズムを開発する外、皺眉筋筋電の計測方法を改良する必要がある。

視線情報から実験参加者の作業状態を検出し、作業状態の遷移を特徴とした推定モデルでは、より正しく戸惑い度と順調度を推定できる結果を得た。今後の課題では、作業状態の遷移に関する指標を定性的に分析し、推定に役立つ指標を選出し、より精度が高いモデルを構築する。また、開発されたモデルを既存の 1 対多遠隔コラボレーションシステムに導入し検証実験を行う予定である。

【参考文献】

- [1] TY Wang, M. Otsuki, & H. Kuzuoka. Evaluating workload in one-to-many remote collaboration. In Proceedings of the 15th International Conference on Pervasive Technologies Related to Assistive Environments, (2022, June), pp. 296-297
- [2] X. Fan, C. Zhao, X. Zhang, H. Luo, & W. Zhang. Assessment of mental workload based on multi-physiological signals. Technology and Health Care (2020), 28(S1), 67-80.
- [3] D Colombini, E Occhipinti, G Molteni, A Grieco, A Pedotti, S Boccardi, C Frigo, and O Menoni. 1985. Posture analysis. Ergonomics 28, 1 (1985), 275–284.
- [4] Brian D Lowe, Patrick G Dempsey, and Evan M Jones. 2019. Ergonomics assessment methods used by ergonomics professionals. Applied ergonomics 81 (2019), 102882.
- [5] Nwakanma, C.I.; Islam, F.B.; Maharani, M.P.; Lee, J.-M.; Kim, D.-S. Detection and Classification of Human Activity for Emergency Response in Smart Factory Shop Floor. Appl. Sci. 2021, 11, 3662.
- [6] Zhongke Gao, Xinmin Wang, Yuxuan Yang, Chaoxu Mu, Qing Cai, Weidong Dang, and Siyang Zuo. 2019. EEG-based spatio-temporal convolutional neural network for driver fatigue evaluation. IEEE transactions on neural networks and learning systems 30, 9 (2019), 2755–2763.
- [7] Jennifer F May and Carryl L Baldwin. 2009. Driver fatigue: The importance of identifying causal factors of fatigue when considering detection and countermeasure technologies. Transportation research part F: traffic psychology and behaviour 12, 3 (2009), 218–224.
- [8] Myounghoon Jeon and Bruce N Walker. 2011. What to detect? Analyzing factor structures of affect in driving contexts for an emotion detection and regulation system. In Proceedings of the Human Factors and Ergonomics Society Annual Meeting, Vol. 55. SAGE Publications Sage CA: Los Angeles, CA, 1889–1893.
- [9] Ebrahim Babaei, Namrata Srivastava, Joshua Newn, Qiushi Zhou, Tilman Dingler, and Eduardo Velloso. 2020. Faces of focus: A study on the facial cues of attentional states. In Proceedings of the 2020 CHI conference on human factors in computing systems. 1–13.

- [10] Gabriele Rescio, Andrea Manni, Marianna Ciccarelli, Alessandra Papetti, Andrea Caroppo, and Alessandro Leone. 2024. A Deep Learning-Based Platform for Workers' Stress Detection Using Minimally Intrusive Multisensory Devices. *Sensors* 24, 3 (2024), 947.
- [11] Oskar Palinko, Andrew L Kun, Alexander Shyrovkov, and Peter Heeman. 2010. Estimating cognitive load using remote eye tracking in a driving simulator. In *Proceedings of the 2010 symposium on eye-tracking research & applications*. 141–144.
- [12] Agnese Brunzini, Fabio Grandi, Margherita Peruzzini, and Marcello Pellicciari. 2023. An integrated methodology for the assessment of stress and mental work-load applied on virtual training. *International Journal of Computer Integrated Manufacturing* (2023), 1–19.
- [13] Sidney K D'Mello and Arthur C Graesser. 2014. Confusion. In *International handbook of emotions in education*. Routledge, 289–310.
- [14] Sidney D'Mello, Blair Lehman, Reinhard Pekrun, and Art Graesser. 2014. Confusion can be beneficial for learning. *Learning and Instruction* 29 (2014), 153–170.
- [15] Jonathan Z Bakdash and Laura R Marusich. 2017. Repeated measures correlation. *Frontiers in psychology* 8 (2017), 456.
- [16] Francis T Durso, Kaitlin M Geldbach, and Paul Corballis. 2012. Detecting confusion using facial electromyography. *Human factors* 54, 1 (2012), 60–69.

〈発 表 資 料〉

題 名	掲載誌・学会名等	発表年月
Exploring relationship between EMG, confusion and smoothness of work progress in assembly tasks	AHs '24: Proceedings of the Augmented Humans International Conference 2024	2024 年 4 月