

令和 6 年 6 月 18 日現在

機関番号：34315
研究種目：研究活動スタート支援
研究期間：2022～2023
課題番号：22K21316
研究課題名（和文）モダリティアテンションネットを用いたマルチモダル情動データによるうつ病診断
研究課題名（英文）Depression Diagnosis Using Multimodal Emotion Data with Modality Attention Networks
研究代表者
劉 家慶（Liu, Jiaqing）
立命館大学・情報理工学部・助教
研究者番号：20948343
交付決定額（研究期間全体）：（直接経費） 2,200,000円

研究成果の概要（和文）：本研究は、表情、音声、および歩行といったマルチモーダル情報を統合することにより、うつ病の診断精度を向上させることを目的とした。特に、モダリティアテンションネットワークを活用して、各モダリティからの情報を効果的に融合し、うつ状態のより正確な認識を実現する新しい手法を提案する。この研究成果は、臨床診断の精度向上に寄与し、うつ病治療の質を高めることで社会的な影響も期待される。

研究成果の学術的意義や社会的意義

本研究は、モダリティアテンションネットワークを用いて、うつ病の正確な診断と早期発見を可能にする新しいアプローチを提案した。この方法は、複数のモダリティからの情報を統合し、診断プロセスの信頼性と精度を向上させる。この研究の成果は、医療診断技術の進展に寄与し、うつ病治療の効率化と患者の生活質の改善につながる社会的意義も持つ。

研究成果の概要（英文）：This study aimed to enhance the accuracy of depression diagnosis by integrating multimodal information such as facial expressions, voice, and gait. Utilizing the Modality Attention Network, a new method was proposed to effectively fuse information from various modalities, achieving more accurate recognition of depressive states. The findings contribute to improving the precision of clinical diagnoses and are expected to have a significant social impact by enhancing the quality of depression treatment.

研究分野：人工知能

キーワード：マルチモダリティAI うつ状態自動認識 情動データベース 動的歩行

1. 研究開始当初の背景

診療現場におけるうつ病の診断は、ストレスが多い現代社会において国内外で深刻な問題となっている。正確な診断に基づく「うつ状態」の早期発見が重要であるとされており、一般的な問診のみならず、患者の表情や音声などの複数の観測情報（モダリティ）を利用したうつ病診断が取り組まれている。最近では、深層学習を用いたうつ状態の認識を支援する研究が増えているが、問診や表情、音声など単一モダリティからの情報を基にした学習が中心であり、その認識精度はまだ不十分である。先行研究では、表情と音声の複数モダリティを融合させた学習方法によりうつ状態の認識精度を向上させたものの、単一モダリティをそれぞれ入力とした複数のネットワークを最終層で結合する形式のため、得られた特徴の解析が不十分であった。また、うつ状態では歩行パターンも日常と異なることが観測されるため、歩行特徴はうつ状態の診断において重要な情報とされ、臨床現場でも注目されている。事前の研究では、歩行のみで感情状態を識別するアプローチが採用され、気分に応じて歩行動作に相違が存在することが示されている。

2. 研究の目的

本研究は、マルチモダリティ情報に基づくうつ状態認識と診断支援を目的とする。先行研究では表情と音声のみでの解析が行われていたが、新たなモダリティとして「歩行」を導入することで、より高精度なうつ状態の推定を目指す。また、モダリティ間の融合がうつ状態をより高精度に認識する可能性を示す知見が先行研究から得られている。うつ状態をより正確に認識するためには、各モダリティがどの程度影響するかを検討する必要がある。このため、モダリティ間の重要性を自動的に学習し、特徴を選択的に融合可能な新たな深層学習手法を提案する。

3. 研究の方法

研究の方法としては、以下の3つのアプローチが採用されている。第一に、キーポイントアテンション GCN を用いた関節動作の重要度の可視化を行う。第二に、疑似ポイントを追加した人体グラフ構造の有効性を明らかにする。第三に、複数のモダリティから有効な特徴を選択し統合する手法を確立する。以下に各項目について詳述する。

(1) キーポイントアテンション GCN を用いた関節動作の重要度可視化

既存の GCN 法では全てのキーポイントを一律に（同じ重みで）取り扱っているため、歩行による感情認識の精度が低いという問題が存在する。本研究では、図 2 に示すようなキーポイントアテンション GCN を用いて認識精度の向上を図る。図 1(b)のように、動作に関連するキーポイント（例えば左膝）を自動的に特定し、それらのキーポイントに大きな重み（赤色が大きな重みを示す）を付加することにより、機微な動作に注目することが可能となる。この方法により、機微な動作に関連するキーポイントを中心に解析を行い、機微な動作に対しても高精度に認識することが可能である。

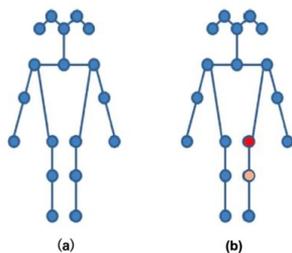


図 1 : (a) 従来の GCN における各キーポイント（関節点）の一律な重み分布；
 (b) 提案手法による各キーポイント（関節点）の重み分布
 （動作に関連するキーポイント（左膝）が高い重みが付与されている（赤色）

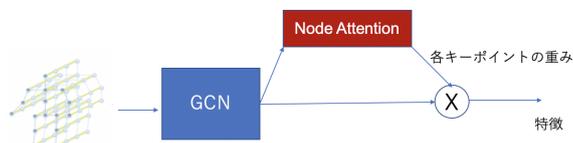


図 2 : キーポイント (node) アテンショングラフ畳み込みネットワーク

(2) 疑似ポイントを追加した人体グラフ構造の有効性

従来の GCN を用いた歩行または動作認識において、人体のグラフ構造は、図 3(a) または (b) のような自然連結または大域連結構造を持つグラフを用いていたが、体の部位によってキーポイント（ノード）の受容野（特徴抽出領域）が制限されており、その部位に関連する動作の認識精度が低下していた。例として、体中心の点（図 4(a)）と手首の点（図 4(b)）の受容野が異なることが示されている。特に手首の受容野が小さいため、認識精度が低い。本研究では、図 3(c) のような疑似キーポイント（緑の点）を追加した全連結グラフ構造を用いる手法を提案する。図 5 に示すように、疑似キーポイントを追加することで、全てのキーポイント（ノード）が同じ広域な受容野を持ち、すべての体部位の動作を高精度で認識することが可能となる。

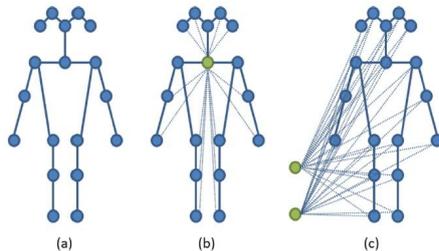


図 3: 人体グラフ構造モデル

- (a) 従来法の natural link model; (b) 従来法の global link model;
(c) 提案法の疑似キーポイント（緑の点）を追加した全連結モデル

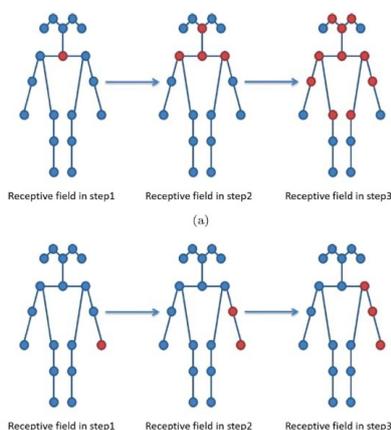


図 4: 従来法におけるキーポイントの受容野（特徴抽出範囲）の変化
(a) 中心関節点; (b) 手首の関節点。

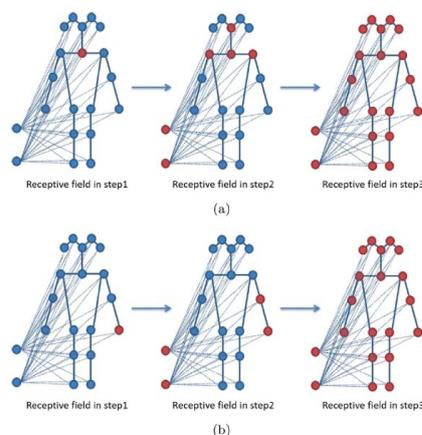


図 5: 提案法におけるキーポイントの受容野（特徴抽出範囲）の変化
(a) 中心関節点; (b) 手首の関節点。

(3) 複数のモダリティから有効な特徴を選択し統合する手法の確立

本研究で、Squeeze-and-Excitation Network (SENet) を拡張発展させたモジュールを用いる。深層学習におけるうつ状態の自動認識において、マルチモダリティ情報の効率的な融合が重要である。各モダリティ情報には相補性が存在し、高い認識精度が期待される一方で、認識に無関係な情報も含まれるため、その取舍選択が重要となる。この問題に対処するために、学習によりモダリティ間の重みを自動的に調整し、認識に重要なモダリティを強調するモダリティアテンションモジュール (図 6) を提案する。このモジュールは、表情と歩行の各モダリティから抽出したそれぞれの特徴マップ ($C \times H \times W$) を Global Average Pooling で ($C \times 1 \times 1$) に圧縮する (Squeeze プロセス) で構成される。この二つの特徴ベクトルを連結し、最後に全連結層 (F_{ex}) で各モダリティに特化したチャンネル重みベクトルを学習し、各特徴ベクトルへ掛け合わせる。これにより、計算されたチャンネル重みベクトルは各モダリティの情報だけでなく、全モダリティ情報を考慮に入れたチャンネルの重みベクトルとなる。また、図 1 に示すように、モダリティアテンションモジュールは層ごとに導入され、得られた各層の特徴を連結し、分類器でうつ状態 (健常、軽症、重症) を鑑別する。

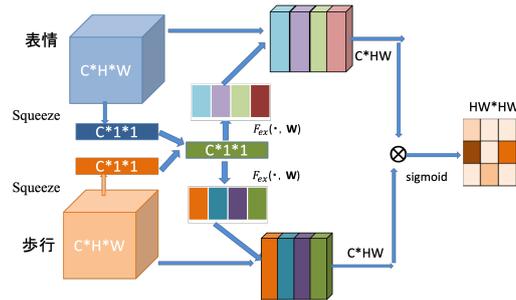


図 6 : モダリティアテンションモジュール (MAM)

4. 研究成果

(1) 関節動作重要度の可視化結果

本研究では、人間の歩行に注目し、時空間グラフニューラルネットワークを用いた歩行からの感情認識を行った。スケルトンベースの動的認識の ST-GCN をベースラインとし、それに対しノード方向、チャンネル方向およびその両方に Attention 機構を導入し、出力の Attention を考慮したモデルを作成することで、認識精度の改善を図った。怒り感情ラベルに対する各 Node (関節点) の重要度 (Attention 重み) の可視化例を図 7 に示す。ノード方向のアプローチが使用したグラフに対してどの程度寄与しているかを確認した。怒りラベルでは、体全体が重要であり、さらに左側の手や足の動きが分類において重要であるということも同時に示され、平常ラベルでは、下半身の動きも重要で、そして特徴的なものとして左右対称な動きが分類に重要であると示された。幸せラベルでは、平常ラベル同様下半身の動きを重要であると示し、特徴的なものとして左膝、左足の動きが重要であることが可視化から示された。そして、悲しみラベルでは下半身の、特に、首から腰回りの動きが重要であり、特徴的なもので左腰、左足の動きを重要であると示された。以上の結果から、歩行という動作において全身の動きが重要であり、特に下半身の動きは共通して重要であることが確認された。

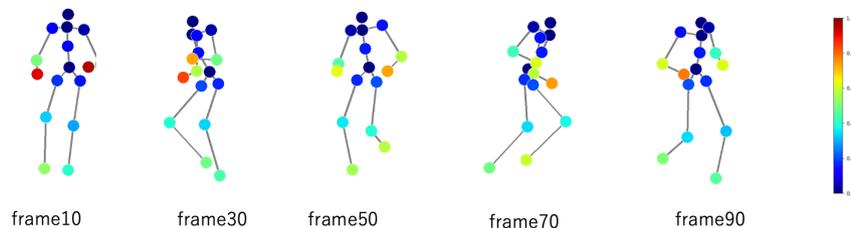


図 7 : 怒り感情に対する各 Node (関節点) の重要度 (Attention 重み) の可視化例

(2) 疑似ポイントを追加した人体グラフ構造の有効性

本研究で提案するキーポイントアテンション GCN 及び疑似ポイントを追加した人体グラフ構造の有効性は、公開データセットを用いた歩行による感情認識の先行研究において既存の state-of-the-art 法に比べ、最も高い認識精度が達成された (表 1)。また、この分野の最新技術として、以下の著名な国際学術誌 Neurocomputing (Impact Factor: 5.779) と日本の画像電子学会誌に掲載された。

表 1 : 提案手法と世界最先端手法との比較 (太い文字は提案法である)

	AP				mAP	Pre	Acc
	H	S	A	N			
ST-TR[44]	0.91	0.54	0.42	0.30	0.5427	0.7903	0.7882
DGNN[18]	0.82	0.37	0.24	0.14	0.3927	0.7893	0.7919
MS-G3D[21]	0.98	0.76	0.61	0.41	0.7088	0.8203	0.8130
G-GCSN[11]	0.98	0.75	0.72	0.46	0.7268	0.8390	0.8296
ST-GCN[9]	0.98	0.76	0.73	0.47	0.7340	0.8399	0.8300
ST-GCN*	0.98	0.75	0.71	0.40	0.7087	0.8344	0.8227
STGCN + PN	0.98	0.76	0.73	0.46	0.7304	0.8514	0.8374
2s-AGCN[17]	0.98	0.70	0.47	0.42	0.6445	0.8214	0.8140
2s-AGCN + PN	0.98	0.71	0.48	0.42	0.6470	0.8237	0.8167
MS-AAGCN[37]	0.98	0.68	0.47	0.40	0.6311	0.8282	0.8190
MS-AAGCN + PN	0.98	0.73	0.56	0.44	0.6760	0.8314	0.8264
CTR-GCN[33]	0.99	0.80	0.70	0.86	0.8417	0.8429	0.8310
CTR-GCN + PN	0.99	0.93	0.78	0.91	0.9063	0.8701	0.8319

(3) 計算機実験による検証

マルチモダリティ情報に基づくうつ状態認識を表 2 にまとめる。表に示すように、提案方法は従来法より高精度であることがわかる。また Top 3 特徴を使うことにより最高の精度を実現していることがわかる。提案法の有効性が示された。

表 2 : 提案手法と世界最先端手法との比較 (太い文字は提案法である)

Method	Emotion	CCC \uparrow	RMSE \downarrow	ACC-4 \uparrow
Liu <i>et al.</i> [8]	Positive	0.060	10.13	0.5
Teng <i>et al.</i> [5]		0.151	9.89	0.55
Sun <i>et al.</i> [9]		0.187	9.76	0.6
Cubemlp [6]		0.231	9.66	0.6
Tensorformer [7]		0.233	9.57	0.6
Ours(Intra fusion)		0.221	9.42	0.6
Liu <i>et al.</i> [8]	Neutral	-0.030	10.64	0.55
Teng <i>et al.</i> [5]		0.136	10.03	0.55
Sun <i>et al.</i> [9]		0.184	9.85	0.5
Cubemlp [6]		0.225	9.78	0.55
Tensorformer [7]		0.219	9.78	0.55
Ours(Intra fusion)		0.231	9.67	0.55
Liu <i>et al.</i> [8]	Negative	-0.050	10.59	0.45
Teng <i>et al.</i> [5]		0.171	9.77	0.55
Sun <i>et al.</i> [9]		0.203	9.47	0.55
Cubemlp [6]		0.225	9.78	0.6
Tensorformer [7]		0.219	9.78	0.6
Ours(Intra fusion)		0.246	9.45	0.6
Baseline	All	0.244	9.61	0.6
Liu <i>et al.</i> [8]		0.070	10.26	0.55
Teng <i>et al.</i> [5]		0.192	9.54	0.55
Sun <i>et al.</i> [9]		0.259	9.45	0.6
Cubemlp [6]		0.331	9.47	0.65
Tensorformer [7]		0.334	9.51	0.65
Ours		0.351	9.41	0.65
Ours(Top 3 features)		0.441	9.36	0.65

5. 主な発表論文等

〔雑誌論文〕 計5件（うち査読付論文 5件/うち国際共著 3件/うちオープンアクセス 0件）

1. 著者名 Sun Hao, Liu Jiaqing, Chen Yen-Wei, Lin Lanfen	4. 巻 91
2. 論文標題 Modality-invariant temporal representation learning for multimodal sentiment classification	5. 発行年 2023年
3. 雑誌名 Information Fusion	6. 最初と最後の頁 504 ~ 514
掲載論文のDOI (デジタルオブジェクト識別子) 10.1016/j.inffus.2022.10.031	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 該当する

1. 著者名 劉家慶, 木下将児, 柴樹榕, 健山智子, 岩本祐太郎, 陳延偉	4. 巻 51
2. 論文標題 Dual Attention 機構を取り入れた時空間グラフ畳み込みネットワークによる歩行からの感情認識	5. 発行年 2022年
3. 雑誌名 画像電子学会誌	6. 最初と最後の頁 309 ~ 317
掲載論文のDOI (デジタルオブジェクト識別子) なし	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 Chai Shurong, Liu Jiaqing, Jain Rahul Kumar, Tateyama Tomoko, Iwamoto Yutaro, Lin Lanfen, Chen Yen-Wei	4. 巻 511
2. 論文標題 A multi-head pseudo nodes based spatial-temporal graph convolutional network for emotion perception from GAIT	5. 発行年 2022年
3. 雑誌名 Neurocomputing	6. 最初と最後の頁 437 ~ 447
掲載論文のDOI (デジタルオブジェクト識別子) 10.1016/j.neucom.2022.09.061	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 該当する

1. 著者名 TENG Shiyu, LIU Jiaqing, HUANG Yue, CHAI Shurong, TATEYAMA Tomoko, HUANG Xinyin, LIN Lanfen, CHEN Yen-Wei	4. 巻 E107.D
2. 論文標題 An Intra- and Inter-Emotion Transformer-Based Fusion Model with Homogeneous and Diverse Constraints Using Multi-Emotional Audiovisual Features for Depression Detection	5. 発行年 2024年
3. 雑誌名 IEICE Transactions on Information and Systems	6. 最初と最後の頁 342 ~ 353
掲載論文のDOI (デジタルオブジェクト識別子) 10.1587/transinf.2023HCP0006	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 該当する

1. 著者名 Chai Shurong, Jain Rahul Kumar, Liu Jiaqing, Teng Shiyu, Tateyama Tomoko, Li Yinhao, Chen Yen-Wei	4. 巻 580
2. 論文標題 A motion-aware and temporal-enhanced Spatial-Temporal Graph Convolutional Network for skeleton-based human action segmentation	5. 発行年 2024年
3. 雑誌名 Neurocomputing	6. 最初と最後の頁 127482 ~ 127482
掲載論文のDOI (デジタルオブジェクト識別子) 10.1016/j.neucom.2024.127482	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

〔学会発表〕 計18件 (うち招待講演 2件 / うち国際学会 14件)

1. 発表者名 Jiaqing LIU
2. 発表標題 Multimodal Deep Learning in Depression Estimation
3. 学会等名 2023 5th International Conference on Intelligent Medicine and Image Processing (IMIP 2023) (招待講演) (国際学会)
4. 発表年 2023年

1. 発表者名 Jiaqing LIU
2. 発表標題 Multimodal Deep Learning in Healthcare
3. 学会等名 International Symposium on Advanced Technologies and Applications in the Internet of Things (招待講演) (国際学会)
4. 発表年 2022年

1. 発表者名 Shurong Chai, Jiaqing Liu, Rahul Kumar Jain, Yinhao Li, Tateyama Tomoko, Yen-Wei Chen
2. 発表標題 A Spatial-Temporal Graph Convolutional Networks-based Approach for the OpenPack Challenge 2022
3. 学会等名 2nd International Workshop on Behavior analysis and Recognition for knowledge Discovery (BiRD 2023) (国際学会)
4. 発表年 2023年

1. 発表者名 Masaya Okano, Jiaqing Liu, Tomoko Tateyama, Yutaro Iwamoto, Yen-Wei Chen
2. 発表標題 Accurate Hand Gesture Recognition Using Color and Depth Images with Modality-invariant Fusion
3. 学会等名 2022 IEEE 11th Global Conference on Consumer Electronics (国際学会)
4. 発表年 2022年

1. 発表者名 Shiyu Teng, Shurong Chai, Jiaqing Liu, Tomoko Tateyama, Xinyin Huang, Yen-Wei Chen
2. 発表標題 A Transformer-based Multimodal Network for Audiovisual Depression Prediction
3. 学会等名 2022 IEEE 11th Global Conference on Consumer Electronics (国際学会)
4. 発表年 2022年

1. 発表者名 Hao Sun, Hongyi Wang, Jiaqing Liu, Yen-Wei Chen, Lanfen Lin,
2. 発表標題 CubeMLP: A MLP-based Model for Multimodal Sentiment Analysis and Depression Estimation
3. 学会等名 n Proceedings of the 30th ACM International Conference on Multimedia (MM ' 22) (国際学会)
4. 発表年 2022年

1. 発表者名 Shurong Chai, Rahul Kumar Jain, Jiaqing Liu, Tomoko Tateyama, Yinhao Li, Yen-Wei Chen
2. 発表標題 A Module Selection-based Approach for Efficient Skeleton Human Action Recognition
3. 学会等名 パターン認識・メディア理解研究会 (PRMU)
4. 発表年 2022年

1. 発表者名 Shiyu Teng, Shurong Chai, Jiaqing Liu, Tateyama Tomoko, Xinyin Huang, Yen-Wei Chen,
2. 発表標題 A transformer based feature-level fusion approach for Audiovisual Depression Prediction
3. 学会等名 パターン認識・メディア理解研究会 (PRMU)
4. 発表年 2022年

1. 発表者名 岡田 一真, 寺田 卓馬, 劉 家慶, 健山 智子, 木村 亮介, 陳 延偉
2. 発表標題 pointnet++に基づく 遺伝子関連研究に向けた3次元顔面形態認識
3. 学会等名 映像情報メディア学会
4. 発表年 2022年

1. 発表者名 Shiyu Teng, Shurong Chai, Jiaqing Liu, Tomoko Tateyama, Lanfen Lin, Yen-Wei Chen
2. 発表標題 Multi-Modal and Multi-Task Depression Detection with Sentiment Assistance
3. 学会等名 IEEE 42nd International Conference on Consumer Electronics (国際学会)
4. 発表年 2024年

1. 発表者名 Kazuma Okada, Takuma Terada, Ryosuke Kimura, Jiaqing Liu, Tomoko Tateyama, Yen-Wei Chen
2. 発表標題 3D Facial Ethnicity Identification Using Point Cloud Deep Learning with Local Area Attention
3. 学会等名 IEEE 42nd International Conference on Consumer Electronics (国際学会)
4. 発表年 2024年

1. 発表者名 Masaya Okano, Jiaqing Liu Tomoko Tateyama Yen-Wei Chen
2. 発表標題 Skeleton-Based Dynamic Hand Gesture Datasets: Recognition and Baseline Evaluation for Deep Learning
3. 学会等名 IEEE 42nd International Conference on Consumer Electronics (国際学会)
4. 発表年 2024年

1. 発表者名 Chengyi Qu, Chaise Ballotti, Daniel De Sousa, Jiaqing Liu
2. 発表標題 Intelligent UAS-Edge-Server Collaboration and Orchestration in Disaster Response Management
3. 学会等名 31th IEEE International Conference on Enabling Technologies(WETICE-2023) (国際学会)
4. 発表年 2023年

1. 発表者名 Liang Lyu, Shurong Chai, Jiaqing Liu, Tomoko Tateyama, Xu Qiao, Yen-Wei Chen
2. 発表標題 A 3D Fusion U-Net with Dual CNN and Transformer Encoders for Lung Airway Segmentation
3. 学会等名 2023 IEEE 12th Global Conference on Consumer Electronics (国際学会)
4. 発表年 2023年

1. 発表者名 Kazuma Okada, Takuma Terada, Ryosuke Kimura, Jiaqing Liu, Tomoko Tateyama, Yen-Wei Chen
2. 発表標題 3D Facial Ethnicity Identification Using PointNet++ with Data Augmentation Based on Farthest Point Sampling
3. 学会等名 2023 IEEE 12th Global Conference on Consumer Electronics (国際学会)
4. 発表年 2023年

1. 発表者名 Jiaqing Liu, Shuqiong Wu, Fumio Okura, Yasushi Makihara, Yasushi Yagi
2. 発表標題 Two-stream Graph Convolutional Networks with Task-specific Loss for Dual-task Gait Analysis,
3. 学会等名 the 45th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC '23) (国際学会)
4. 発表年 2023年

1. 発表者名 Shurong CHAI, Rahul Kumar JAIN, Yin hao Li, Jiaqing Liu, Tomoko Tateyama and Yen-Wei Chen
2. 発表標題 Adaptive Graph Convolutional Networks for Medical Image Segmentation,
3. 学会等名 the 45th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC '23) (国際学会)
4. 発表年 2023年

1. 発表者名 Shiyu Teng, Jiaqing Liu, Shurong Chai, Tomoko Tateyama, Xinyin Huang, Lanfen Lin, Yen-Wei Chen
2. 発表標題 An Intra- and Inter-Emotion Transformer-based Fusion Model with Homogeneous and Diverse Constraints Using Multi-emotional Audiovisual Features for Depression Detection
3. 学会等名 第26回 画像の認識・理解シンポジウム (MIRU2023), 浜松
4. 発表年 2023年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

-

6. 研究組織

氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考

7. 科研費を使用して開催した国際研究集会

〔国際研究集会〕 計0件

8 . 本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関
---------	---------