

デュアルタスクにおける認知機能推定のための Attention機構を用いた関節動作の重要度可視化

野口 智矢^{1,a)} 武 淑瓊^{1,b)} 劉 家慶^{1,c)} 大倉 史生^{1,d)} 槇原 靖^{1,e)} 八木 康史^{1,f)}

概要：認知症は薬やリハビリによって進行を遅らせることができるものの完治はできない。そのため日常的に認知機能を測り、MCI（軽度認知障害）を早期発見することが求められている。この問題に対して、足踏みと計算を同時に行うデュアルタスクにより、認知機能低下の検出を目指す研究が行われている。しかしモデルがブラックボックス化しているため、社会実装や精度改善の障壁となっている。本研究ではブラックボックス解明の検証段階として、足踏み時における関節点の重要度を可視化することにより、モデルの注目領域の理解を目指す。そのために従来のアーキテクチャに対して、チャンネル方向と時空間方向のAttentionを取得するモジュールを追加した、2種類の手法を提案した。さらに、一例ずつの可視化だけでなく、全サンプルの傾向や特徴を掴むための分析を行った。具体的には、全サンプルを複数のグループにクラスタリングし、各グループ毎に各関節の平均重要度を求めた。その結果、認知機能のレベル毎に異なる関節点を注目していることを明らかにした。

キーワード：デュアルタスク、認知症、MCI（軽度認知障害）、ST-GCN（時空間グラフ畳み込みニューラルネットワーク）、MMSE（ミニメンタルステート検査）

1. はじめに

認知症は、脳の病気や障害などの原因で認知機能が低下し、日常生活全般に支障が出てくる状態である。昨今の日本では高齢化が進んでおり、それに伴い認知症の数も増えている。2025年には約700万もの人が認知症の有病者になることが予想されている。

認知症は完全に治療する方法が見つからないが、リハビリや薬を処方することで症状の進行を遅らせることはできる。完全に認知症へ移行する前に、MCI (Mild Cognitive Impairment: 軽度認知障害) と呼ばれる、認知症の初期段階となるステージを挟むことがわかってきている。そのため日常的に検査を行うことによって、MCIの段階で認知障害の早期発見をすることが重要である。

認知機能を測る検査として、MMSE (Mini-Mental State Examination: ミニメンタルステート検査) [1] があり、国際的に広く用いられている。MMSEは認知機能のレベルを30点満点で点数化することにより、「0以上23点以下な

ら認知症疑い」「24以上27点以下ならMCI疑い」「28点以上なら正常」と分類することができる。しかしこの認知機能検査には検査の際に医師等の専門家の立ち合いが必要であったり、問題への対策ができてしまうという点から日常的な検査に向かないという欠点がある。

近年、複数の動作を同時に行うデュアルタスクが認知症の検出に有効であることが報告されており [2] [3] [4]、また日常的な検査に向いているという点から、足踏みと計算問題を組み合わせたデュアルタスクによって認知症検出を試みるシステムが開発された [5] [6]。このシステムは、デュアルタスクによって得られた足踏みデータと計算問題の解答データから、機械学習手法を用いて認知機能の推定を行うものである。しかし機械学習モデルへの入力として与える特徴量を選択する必要があるため、デュアルタスク時の全情報を活用しきれていないという問題点がある。

その問題に対し武らはST-GCN (Spatial temporal graph convolutional networks: 時空間グラフ畳み込みネットワーク) [7] と呼ばれる、骨格情報から人物の動作認識をする機械学習モデルをデュアルタスクシステムに導入した [8]。その結果従来手法と比較して精度の向上に成功したが、モデルがブラックボックス化しているために、解釈性や説明性に欠けるという問題点が残っている。

この問題を解決するために我々は、ブラックボックス解

¹ 大阪大学

a) noguchi@am.sanken.osaka-u.ac.jp

b) wu@am.sanken.osaka-u.ac.jp

c) liu@am.sanken.osaka-u.ac.jp

d) okura@ist.osaka-u.ac.jp

e) makihara@am.sanken.osaka-u.ac.jp

f) yagi@am.sanken.osaka-u.ac.jp

明の検証段階として、ST-GCN モデルに対して関節動作の重要度可視化を行った。

本研究の貢献は以下の2点である。

(1) チャンネル方向の Attention を用いた関節動作の重要度可視化と分析: 従来手法の ST-GCN に対し、チャンネル方向の Attention 機構を導入したネットワークを新たに提案する。認知機能レベル毎に異なる、注目すべきチャンネルを自動的に学習させ、特徴マップに重み付けをすることによって、足踏み動作の重要度可視化を実現した。

さらにワンサンプルの可視化結果だけでなく、複数サンプル間での共通性や傾向を調べるために、クラスタリングを用いた分析を行う。その結果、健常者・MCI 疑いの分類を行う際生成される特徴マップに、明確な差が確認できた。

(2) Attention を用いた手法比較: チャンネル方向の Attention を用いた手法と、時空間方向の Attention を用いた手法の比較を行う。本研究で扱うような小さなデータベースでは、STA-GCN が適切でないことを確認した。

2. 関連研究

2.1 デュアルタスクシステム

異なる2つの動作を同時に行うデュアルタスクをすることは、認知機能が低い人にとってタスクの処理領域の切り替えが難しくなるため、認知症の検出に役立つことが報告されている [2] [3]。そこで本研究グループでは、足踏みと計算問題を組み合わせたデュアルタスクを行わせるシステムも開発された [5]。このシステムは被験者に1. 計算問題 (シングルタスク)、2. 足踏み (シングルタスク)、3. 計算しながら足踏み (デュアルタスク) の3つのタスクを課す。それにより、計算の解答データや足踏み時の関節動作のデータを得ることができる。ここで関節動作データは、各フレーム毎に25個の関節点が検出され、各関節点は3次元座標で表現される。すなわち、3次元座標 × 関節数 (25) × フレーム数で表現される。

得られた解答データと関節動作データから「足踏み平均速度 [steps/sec]」「足踏み速度標準偏差」「膝角度平均 [rad]」「膝角度標準偏差」「認知タスク正答率」「認知タスク回答回数平均」の特徴量を抽出する。これらの特徴量を LightBGM 等の機械学習手法に入力し、MMSE スコアを教師ラベルとして学習することで、被験者の認知機能の推定をすることができる [6]。

2.2 ST-GCN によるデュアルタスクデータからの認知機能推定

前節で記した LightBGM による分類には、関節動作データから「足踏み平均速度 [steps/sec]」等の特徴量を直接求める必要がある。しかしこの方法では、頭部の左右の揺れ等の足踏みの様子全体を捉えきれないという問題点があった。そこで武らはデュアルタスクシステムによって得られ

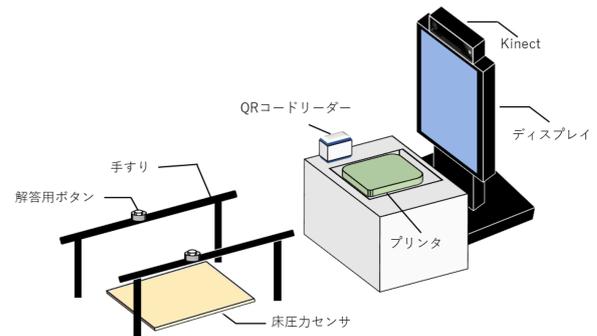


図1 デュアルタスクシステムの外観

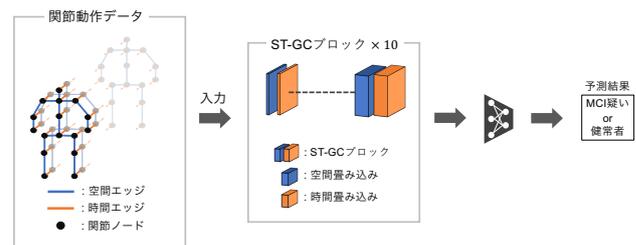


図2 ST-GCN の概要図

た25個の関節点の3次元座標データに対して、ST-GCN (Spatial Temporal Graph Convolutional Networks: 時空間グラフ畳み込みネットワーク) [7] を適用することを提案した。ST-GCN とは図2に示すようなグラフ構造として表現された関節動作データから、グラフ畳み込み [9] をすることによって空間パターンと時間パターン両方を学習する動作認識モデルである。さらに武らは、医学分野における診断方法の評価指標として用いられている「感度 + 特異度」を直接最適化する損失関数を設計した。その結果認知症の閾値となる MMSE スコアが24以上・24未満の分類において「感度 + 特異度」が1.64、MCIの閾値となる MMSE が28以上・28未満の分類においては1.76という精度を達成した。

しかしモデルがブラックボックス化しており、一連の関節動作の中で、どこを注目して分類を行なっているのか分からないという問題点が残っている。

3. 提案手法

本研究では、ST-GCN モデルに対する関節動作の重要度可視化のために、Attention を用いた2つの手法を提案する。具体的には、畳み込み演算におけるチャンネル方向の Attention を用いた可視化手法と、時空間方向の Attention を用いた可視化手法を試みた。

3.1 手法 (1): チャンネル Attention を用いた関節動作の重要度可視化

本手法では、ニューラルネットワークから生成される特

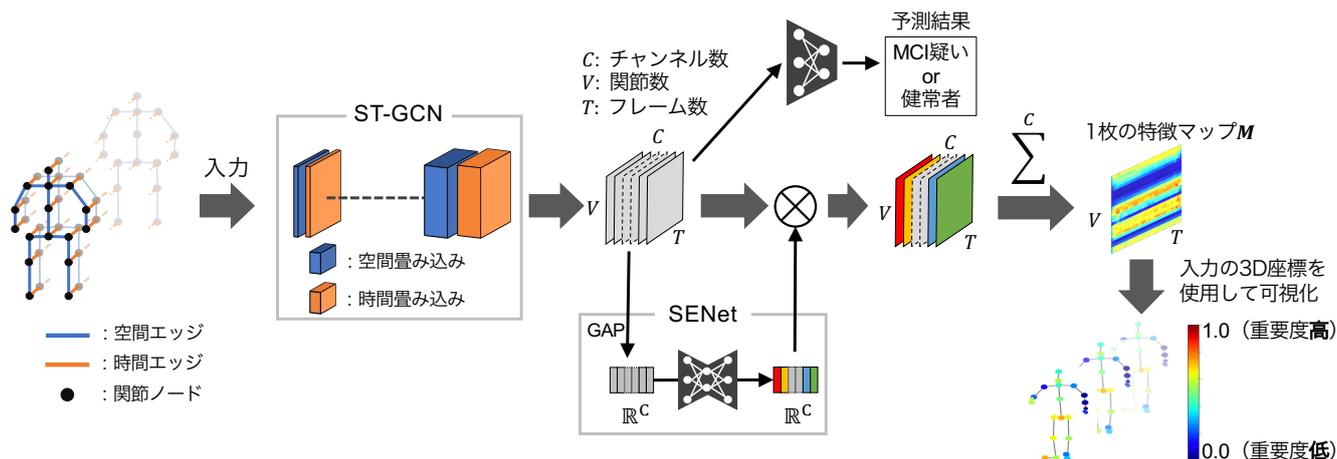


図 3 手法 (1) の概要図

微マップを使用することにより、モデルの判断根拠可視化を試みる。モデルの最終層の特徴マップをもとに顕著性マップを推論する手法 [10] が広く知られているため、本研究ではこのアイデアを踏襲する。

ST-GCN はグラフ畳み込み [9] を伴うため、複数のチャンネル分の特徴マップが生成される（本研究で扱うモデルは最終的に 256 のチャンネルを生成する）。ここで問題になるのが、複数チャンネル分ある特徴マップの中で「どのように複数のチャンネルを結合するか」という点である。単純に全ての特徴マップを足し合わせた場合、重要度の低い情報と重要度の高い情報が打ち消し合ってしまう。そこで本手法では複数チャンネルに対し重要度となる重みを求め、チャンネルの重みと特徴マップの線型結合をとることにより、図 3 のように 1 枚の特徴マップ M を取得する。

チャンネル重要度（チャンネルの重み）の求め方を説明する。チャンネルの重要度の求め方として、様々なアプローチが存在する [10] が、本研究ではチャンネルに対する Attention 機構を持つ、SENet (Squeeze-and-Excitation Networks) [11] を用いる。

SENet は計算量やパラメータ数の増加を抑えることができ、畳み込みを行う様々なモデルに対して追加・精度向上が見込めることから、ST-GCN にも有効であると考えた。本手法では図 3 に示すように、従来の ST-GCN に対してチャンネル重要度を抽出するモジュールを組み込んだネットワークを構築する。

このモジュールでは最初に特徴マップ $\mathbf{U} \in \mathbb{R}^{T \times V \times C}$ ($\mathbf{U} = [\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_C]$) をプーリング層に通すことでグローバルな時空間情報を収集する。 T は時間方向のフレーム数、 V はフレーム毎の関節数を示す。

$$z_c = \frac{1}{T \times V} \sum_{i=1}^H \sum_{j=1}^W u_c(i, j) \quad (1)$$

ここで z_c は合計 C 枚ある内の c 番目の特徴マップ $\mathbf{u}_c \in \mathbb{R}^{T \times V}$ の統計量である。このように各チャンネル

の統計量 $\mathbf{z} \in \mathbb{R}^C$ を求めていく。次に 2 つの全結合層と Tanh 関数・ReLU 関数によってどのチャンネルを残すかを決定するマスク $\mathbf{s} \in \mathbb{R}^C$ を 0 以上 1 以下の範囲で決める。

$$\mathbf{s} = \delta(\tanh(\text{BN}(\mathbf{W}_2 \delta(\mathbf{W}_1 \mathbf{z})))) \quad (2)$$

ここで δ は ReLU 関数、 BN はバッチ正規化、 $\mathbf{W}_1 \in \mathbb{R}^{\frac{C}{16} \times C}$ と $\mathbf{W}_2 \in \mathbb{R}^{C \times \frac{C}{16}}$ は全結合層 (FC 層) を指す。

その結果チャンネル毎の重要度マスク \mathbf{s} を求めることができる。最終的には図 3 に示すようにチャンネル重要度のマスクを $\mathbf{u}'_c = s_c \times \mathbf{u}_c$ と元の特徴マップに掛けることで、チャンネルの重要度によって重み付けされた特徴マップ $\mathbf{U}' \in \mathbb{R}^{T \times V \times C}$ を得ることができる。

$$\mathbf{M} = \sum_{c=1}^C \mathbf{u}'_c \quad (3)$$

得られた特徴マップ \mathbf{U}' に対し、式 (3) のようにチャンネル方向に足し合わせることで、図 3 のような 1 枚の特徴マップ $M \in \mathbb{R}^{T \times V}$ を得る。

上記で得られた 1 枚の特徴マップ $M \in \mathbb{R}^{T \times V}$ は時間方向の次元と関節の次元から構成される。そのためこの特徴マップを使用すれば、あるフレームにおけるある関節点の重要度を一意に求めることができる。入力として与えた関節動作データを三次元上にプロットし、重要度をカラーで表現することによって、動画として関節動作の重要度可視化をすることができる。

3.2 手法 (2): 時空間 Attention を用いた関節動作の重要度可視化

本手法では従来の ST-GCN に対して、時空間方向の注目領域を学習するような Attention 機構を追加することで、その注目領域を用いた関節動作の重要度可視化を行う。画像認識分野では Attention を組み込んだモデルが多数存在するが、本手法では ST-GCN に Attention 機構を追加した STA-GCN (Spatial Temporal Attention Graph

Convolutional Networks) [12] を採用する. この手法は「動作認識において注目すべき関節の度合い」と「相関性の高い関節の関係性」を学習するグラフ畳み込みニューラルネットワークである.

まず本手法の概要について説明する. 図4に示すように本手法ではベースとなる従来の ST-GCN の内, 前半の ST-GC ブロックの4層を”Feature extractor”に後半の ST-GC ブロックの6層を”Perception branch”に分割する. さらに”Attention branch”と呼ばれるモジュールを新たに追加する. これは Feature extractor の出力 \mathbf{X}_{out}^{FE} をもとに注目領域を取得する [13]. この注目領域は, フレーム毎の重要な関節を表現する”Attention node”と関節同士の重要な関係性を表現する”Attention edge”に分けられる.

Attention node は「関節 × フレーム」の2次元マップで表され, 要素の値が大きいほど「あるフレームにおける, ある関節の重要度が高い」と解釈できる. この Attention node は以下のように Feature extractor の出力と乗算することで, 重要な関節の時空間情報を強調することができる.

$$\mathbf{X}'_{out} = M(\mathbf{X}_{out}^{FE}) \cdot \mathbf{X}_{out}^{FE} \quad (4)$$

ここで $M(\mathbf{X}_{out}^{FE})$ は Attention node を指し, この出力 \mathbf{X}'_{out} は Perception branch への入力となる.

Attention edge は「関節 × 関節」の隣接行列で表現され, Perception branch へ直接流れていく. ここで従来の骨格パターンに対するグラフ畳み込み結果 \mathbf{M}_q^{skel} に加えて, Attention edge に対するグラフ畳み込み

$$\mathbf{X}_{out}^{att} = \sum_{k=1}^K \hat{\mathbf{A}}_k^{att} \mathbf{X}_{in} \mathbf{W}_k^{att} \quad (5)$$

を行うことで, 関節同士の重要な関係性を取り込んだ特徴量

$$\mathbf{X}_{out} = \mathbf{X}_{out}^{skel} + \mathbf{X}_{out}^{att} \quad (6)$$

を取得することができる. (5) 式における K は入力の1サンプルから得られる Attention edge の数, $\hat{\mathbf{A}}_k^{att}$ は Attention edge を正規化した隣接行列, \mathbf{W}_k^{att} は重み行列を表す. このとき Attention edge のある要素の値が大きいほど, そのエッジがたが関節点同士の関係性が大きいと解釈することができる.

学習の際には図4にあるように, Attention branch の学習誤差 L_{att} と Perception branch の学習誤差 L_{per} の和に対してバックプロパゲーションすることによって, ネットワーク全体として End-to-end で学習することができる. なお推論時には Perception branch の結果を用いる.

可視化の際には, Attention node と Attention edge の両方を用いる. Attention node は「関節 × フレーム」の2次元マップであることから, あるフレームにおけるある関節点の重要度を一意に求めることができる. そのため, 手法(1)と同様に可視化することができる. Attention edge は

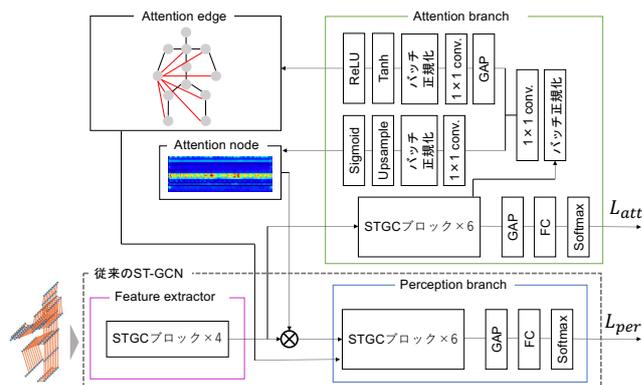


図4 注目領域を取得するために Attention 機構を追加したネットワーク [12]

「関節 × 関節」の隣接行列で表現されることから, 本手法では最大 $25 \times 25 = 625$ 本のエッジが得られる可能性がある. しかしそのような多数の Attention edge を可視化すると非常にみづらくなってしまふ. そのため本手法では大きな値を持つ上位30本を, 関節同士を結ぶようにエッジを可視化することにする.

4. 実験

4.1 実験設定

本研究の実験では, MMSE スコアによって MCI 疑いか健康者かの分類問題を解くモデルを扱う.

本研究で扱うデータベースは, 6つの高齢者施設に設置されたデュアルタスクシステムから収集されたデータである. 収集の結果, 171名の被験者と3271のサンプルが得られた. なお全ての被験者に対してインフォームドコンセントを取得した.

可視化に使用するモデルの学習時には, 層化4分割交差検証によって MMSE スコアの比率を保たせるようにする. その上訓練データとテストデータには同一の体験者のデータが含まれないようにする. 4回の推論のうち, 最も精度の良かったモデルを可視化に使用する. これは精度の良いモデルの方が, データの特徴を的確に捉えていることが期待されるためである.

4.2 手法(1)における関節動作重要度の可視化結果

提案手法の可視化結果を図5と図6に示す. 図5・図6はそれぞれ, モデルが健康者・MCI 疑いと予測し, 教師ラベルも健康者・MCI 疑いとなる例である.

図5からは, 上半身の重要度が常に低く, 左足の重要度が高くなっていることが分かる. 図6も図5と同様に, 上半身の重要度が常に低いことが分かる. しかし右足と左足両方の重要度が高い点が図5と異なる.

4.3 手法(2)における関節動作重要度の可視化結果

手法(2)の可視化結果を図7と図8に示す. 図7と図8

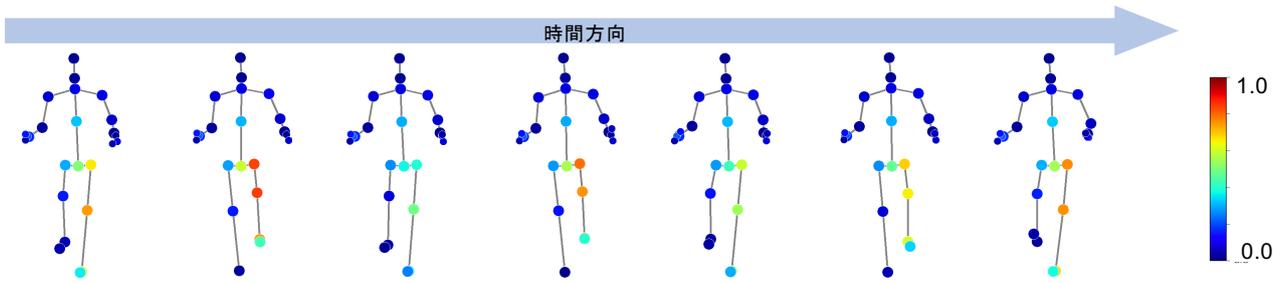


図 5 健常者サンプルの可視化結果

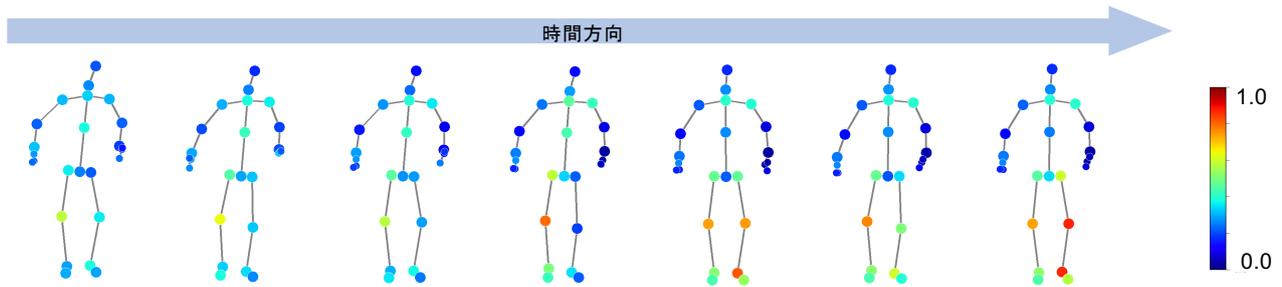


図 6 MCI 疑いサンプルの可視化結果

で可視化をしているサンプルは、図 6 と図 5 で用いているサンプルと同一である。可視化を行うモデルは前セクションと同様に、最も精度の良いモデルを採用した。

図 7 にあるように、このサンプルでは背骨中央部の注目度が常に高いことが分かる。また前半 4 フレームで左腰・後半 3 フレームで右腰の重要度が高くなっていることから、時間とともに注目領域が遷移していると解釈できる。さらに Attention edge を表す赤いエッジが背骨中央部に集中していることより、この関節と他関節の相関性が高いと見ていいることが分かる。それに対して図 8 からは右下半身・頭部・左手周辺の重要度が常に高く、左手や右足に Attention edge が集中していることが分かる。

4.4 手法 (1) の可視化結果分析

前節では健常者と MCI 疑いそれぞれ 1 つのサンプルに対し、どのような重要度が得られるのかを確認した。本節では複数サンプルから生成された特徴マップの中で、「健常者のグループ」と「MCI 疑いのグループ」の間に明確な違いや共通性がないか分析する。なお分析対象となるデータは前節と同様に、MCI 疑い 411 サンプルと健常者 649 サンプルを合わせた 1060 サンプルである。

モデルが生成する特徴マップには多様性があり、1060 サンプル全体に対して一括で有用な統計量を求めるのは困難である。そこで本節では図 9 に示すように、モデルが予測した健常者・MCI 疑いの各グループに対し、k-means 法を用いて 5 つのクラスタに分割する。その後各クラスタにある特徴マップの平均を取ることで、10 枚の代表的な特徴マップ取得する。最終的に、2 次元である特徴マップの時

間方向に対しても平均を取ることで、各クラスタの関節重要度を得ることができる。

図 10 に健常者グループにおける、各クラスタの関節重要度を示す。クラスタ 1 からクラスタ 5 全てに共通して左足の重要度が高いことが分かる。このことから、モデルが健常者と予測するときは左足を注目する傾向があると解釈できる。

図 11 に MCI 疑いグループにおける、各クラスタの関節重要度を示す。図 10 と比較すると、重要な関節点の傾向が健常者グループと異なることがわかる。具体的には、左足だけでなく右足の重要度が高いクラスタや、左肩にも注目しているクラスタが存在する。

このことから健常者グループでは異なるクラスタ間においても、同じ関節点の重要度が高くなる傾向があることが分かる。それに対し MCI 疑いグループでは、異なるクラスタ間において共通性はあるものの異なる関節点の重要度が高くなっていることが分かった。

以上より本分析に用いたモデルにおいては、健常者と予測したグループと MCI 疑いと予測したグループの間で、生成する特徴マップに明確な差があることが確認できた。

4.5 手法 (2) の可視化結果分析

「関節動作重要度の可視化結果」のセクションでは一つのサンプルを入れたときに得られる、関節の重要度を示す「Attention node」を可視化し、どの関節点重要化を見た。本セクションでは得られた重要度が有用なものかどうかの検証を行う。Attention node は図 4 に示すようにシグモイド関数の出力であるため、0 から 1 の範囲の値を取る。

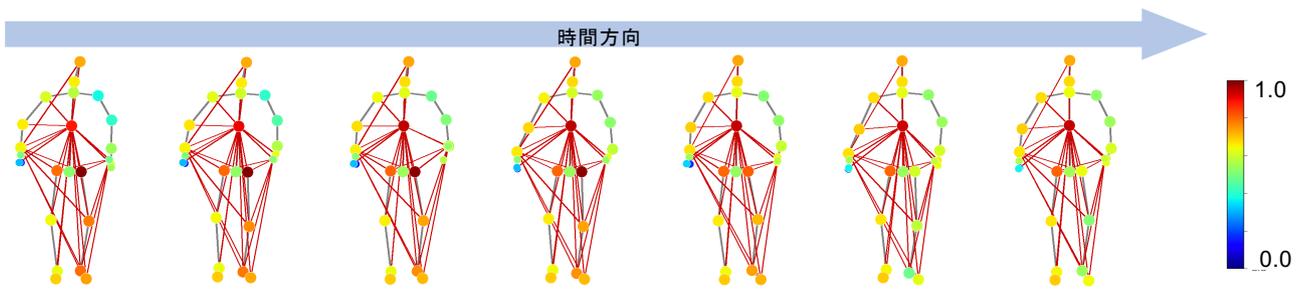


図 7 MCI 疑いのサンプルを手法 (2) で可視化した結果

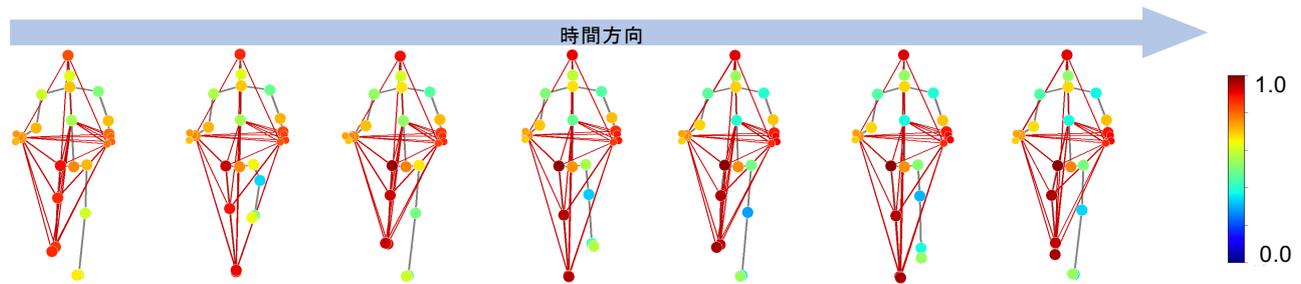


図 8 健常者のサンプルを手法 (2) で可視化した結果

このとき 0 に近ければ重要度は低く、1 に近ければ重要度は高いと解釈できる。手法 (2) ではこの Attention node を可視化するとき、ミニマックス正規化をかけてから色の決定をしていく。ここで問題になるのは Attention node の最小値が 0.1 で最大値が 0.9 のときと最小値が 0.499 で最大値が 0.501 のときで、可視化した際に色のばらつきに違いがなくなってしまうという点である。

まず図 7 と図 8 で可視化した、MCI 疑いと健常者のサンプルの Attention node の範囲を調べる。このとき Attention node は「関節 × フレーム」の 2 次元マップであり、この中から最大値と最小値を 1 つずつ取得する。結果として図 7 のサンプルは最大値 0.5164 最小値 0.4750、図 8 のサンプルは最大値 0.5211 最小値 0.5016 であった。このことから、この 2 サンプルの Attention node はほとんどの値の大きさに違いがないことが分かった。次に全体としてどのような Attention node の値を取るのか調べるために全 1060 サンプルを入力する。その結果全体の最大値が 0.5311 で最小値が 0.4402 であったことから、Attention node の値の範囲が高々 0.1 にも満たないことが分かった。

以上より手法 (2) で可視化する Attention node は、0.5 付近に強い偏りを持つことがいえる。

4.6 考察

まず手法 (1) について考察を行う。手法 (1) では、モデルが MCI 疑いと健常者に分類する際に生成する特徴マップに明確な差がみられた。このことからモデルが陽性と陰性に分類する際に、それぞれ異なる関節点を注目していることが考えられる。

表 1 パラメータ数の比較

手法	パラメータ数
ST-GCN	3,083,926
手法 (1)	3,155,094
手法 (2)	6,386,390

本研究では k-means 法を用いて MCI 疑いグループと健常者グループをさらに分割し、その中で平均特徴マップを作成した。しかしこれは足踏みの周期や位相等の時間情報を無視したものであるため、これらの情報を考慮した解析を行うことは今後の課題である。また k-means 法以外のクラスタリング手法を用いることや、異なるデータセット・モデルの可視化結果と比較を行うことは、今回とは違う知見が得られることが期待される。

次に手法 (2) について考察を行う。手法 (2) の実験において、関節の重要度が取得できなかった原因として、今回導入した Attention branch(図 4) が 60 の動作クラスを認識するために設計されたものであるという点が考えられる。一般的な動作認識は「投げる」「蹴る」「ジャンプする」といった、人間の目で見ても分かりやすく各クラス間の関節動作の差が大きい。それに対して本研究で扱うデュアルタスクは全サンプルが足踏みをしており、人間の目では把握できないほど陽性陰性間の関節動作の差が小さい。さらにパラメータ数が ST-GCN・手法 (1) と比べて大きい (表 1 参照) ことも、データ数が少ない本研究においてネガティブに働いたことが考えられる。このことから注目領域を取得する部分においてデュアルタスクの特徴を抽出しきれなかったことが、関節の重要度を取得できなかった原因として考えられる。

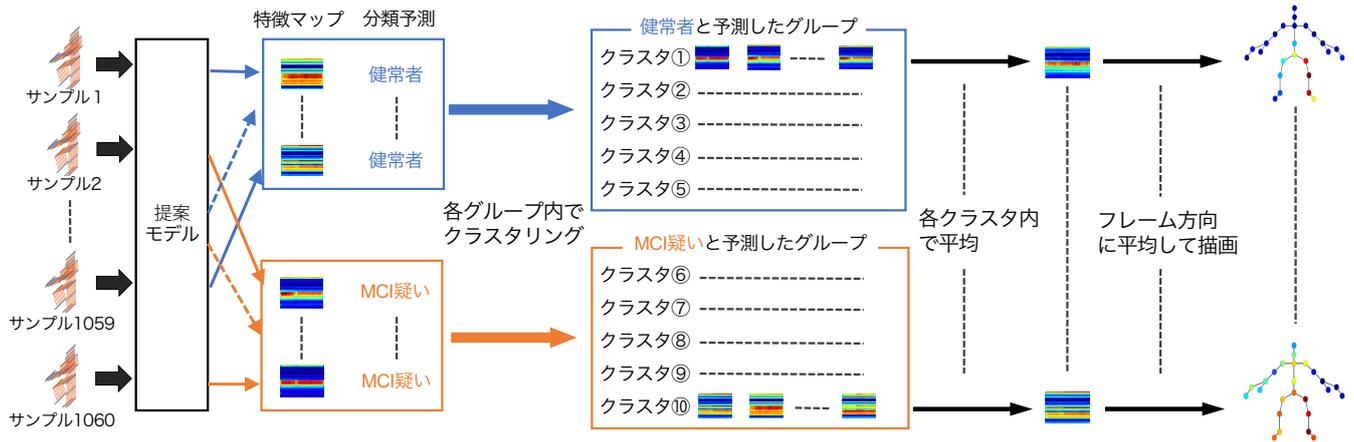


図 9 可視化分析の流れ

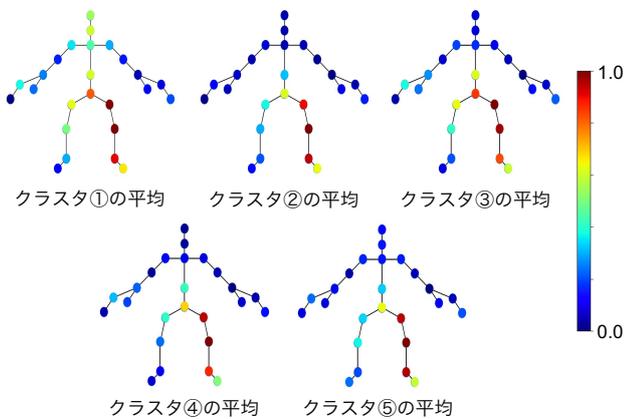


図 10 健常者グループにおける各クラスターの関節重要度

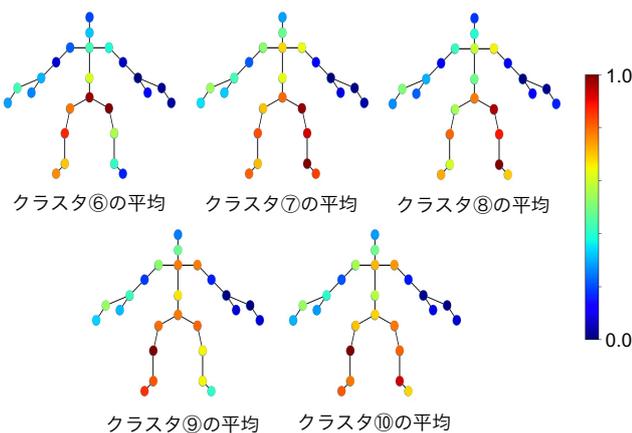


図 11 MCI 疑いグループにおける各クラスターの関節重要度

5. おわりに

本研究では認知機能推定モデルのブラックボックス解明の検証段階として、関節動作の重要度可視化を行った。そのための手法としてまず、チャンネルの重要度によって重み付けされた最終層の特徴マップの可視化を試みた。その結果、MCI 疑い・健常者の分類を行う際生成される特徴

マップに明確な差がみられた。時空間 Attention を用いた手法によって関節動作の注目領域可視化にも試みたが、得られた結果に強い偏りが生じてしまい、有用な情報を得ることができなかった。

チャンネル重要度を用いた手法の今後の課題として、最終層だけでなく中間層の可視化が挙げられる。具体的には層ごとの特徴マップの遷移を可視化することによって、モデルがどの関節動作を見ているのかがより鮮明になる。このことはブラックボックス解明の手がかりとなることが期待される。時空間 Attention を用いた手法の課題として、注目領域を取得するモジュールのデザインが挙げられる。本研究で用いたモジュールは 60 の動作認識用に設計されたものであり、デュアルタスクを利用したデータ専用には設計されていない。このモジュールの修正は認知機能推定モデルのブラックボックス解明だけでなく、モデルの精度向上にも貢献すると考えられる。

今後はこの研究をもとに、認知機能推定モデルの精度と判断根拠性の向上に試みる予定である。

謝辞

本研究の一部は、AMED の課題番号 JP21uk1024001 支援を受けたものである。

参考文献

- [1] Folstein, M. F., Folstein, S. E. and McHugh, P. R.: "Mini-mental state": a practical method for grading the cognitive state of patients for the clinician, *Journal of Psychiatric Research*, Vol. 12, No. 3, pp. 189–198 (1975).
- [2] Åhman, H. B., Cedervall, Y., Kilander, L., Giedraitis, V., Berghlund, L., McKee, K. J., Rosendahl, E., Ingelsson, M. and Åberg, A. C.: Dual-task tests discriminate between dementia, mild cognitive impairment, subjective cognitive impairment, and healthy controls—a cross-sectional cohort study, *BMC Geriatrics*, Vol. 20, No. 1, pp. 1–10 (2020).
- [3] Mancioffi, G., Fiorini, L., Rovini, E., Zeghari, R., Gros, A., Manera, V., Robert, P. and Cavallo, F.: Innovative

- motor and cognitive dual-task approaches combining upper and lower limbs may improve dementia early detection, *Scientific Reports*, Vol. 11, No. 1, pp. 1–9 (2021).
- [4] Smith, E., Cusack, T., Cunningham, C. and Blake, C.: The influence of a cognitive dual task on the gait parameters of healthy older adults: a systematic review and meta-analysis, *Journal of Aging and Physical Activity*, Vol. 25, No. 4, pp. 671–686 (2017).
- [5] Okura, F., Mitsugami, I., Niwa, M., Aoki, K., Zhou, C. and Yagi, Y.: Automatic collection of dual-task human behavior for analysis of cognitive function, *ITE Transactions on Media Technology and Applications*, Vol. 6, No. 2, pp. 138–150 (2018).
- [6] Matsuura, T., Sakashita, K., Grushnikov, A., Okura, F., Mitsugami, I. and Yagi, Y.: Statistical analysis of dual-task gait characteristics for cognitive score estimation, *Scientific Reports*, Vol. 9, No. 1, pp. 1–12 (2019).
- [7] Yan, S., Xiong, Y. and Lin, D.: Spatial temporal graph convolutional networks for skeleton-based action recognition, *Thirty-second AAAI Conference on Artificial Intelligence* (2018).
- [8] Wu, S., Okura, F., Makihara, Y., Aoki, K., Niwa, M. and Yagi, Y.: Early detection of low cognitive scores from dual-task performance data using a spatio-temporal graph convolutional neural network, *2021 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society*, pp. 1895–1901 (2021).
- [9] Kipf, T. N. and Welling, M.: Semi-supervised classification with graph convolutional networks, *arXiv preprint arXiv:1609.02907* (2016).
- [10] Selvaraju, R. R. et al.: Grad-cam: Visual explanations from deep networks via gradient-based localization, *Proceedings of the IEEE ICCV*, pp. 618–626 (2017).
- [11] Hu, J., Shen, L. and Sun, G.: Squeeze-and-excitation networks, *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 7132–7141 (2018).
- [12] Shiraki, K., Hirakawa, T., Yamashita, T. and Fujiyoshi, H.: Spatial temporal attention graph convolutional networks with mechanics-stream for skeleton-based action recognition, *Proceedings of the Asian Conference on Computer Vision* (2020).
- [13] Fukui, H., Hirakawa, T., Yamashita, T. and Fujiyoshi, H.: Attention branch network: Learning of attention mechanism for visual explanation, *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 10705–10714 (2019).