

Title	室内環境における準ミリ波レーダーを使用した姿勢推定システムの研究
Author(s)	崎間, 寛哉
Citation	
Issue Date	2026-03
Type	Thesis or Dissertation
Text version	author
URL	<a href="https://hdl.handle.net/10119/20550">https://hdl.handle.net/10119/20550</a>
Rights	
Description	Supervisor:リム 勇仁, 先端科学技術研究科, 修士(情報科学)

修士論文

室内環境における準ミリ波レーダーを使用した姿勢推定システムの研究

崎間 寛哉

主指導教員 リム 勇仁

北陸先端科学技術大学院大学  
先端科学技術研究科  
(情報科学)

2026年3月

## Abstract

This study investigates a privacy-conscious, non-contact posture estimation system for indoor living environments, aiming to estimate human postural activities without relying on visual sensing. Although camera-based pose estimation methods have achieved high accuracy in recent years, their continuous image acquisition raises significant privacy concerns, which limits their practical deployment in residential and caregiving environments. In contrast, 24 GHz quasi-millimeter-wave radar is robust to low-light conditions and occlusion, and does not capture visual appearance information, making it advantageous from the perspective of privacy protection. However, radar-based observations suffer from low spatial resolution and are highly sensitive to reflections, noise, and environmental dependency, making stable posture estimation using signals from a single sensor challenging.

To address these limitations, this study constructs a learning framework that integrates time-series observations from multiple quasi-millimeter-wave radar sensors by jointly considering spatial relationships among sensors and temporal transitions of human actions. Each radar sensor is treated as a node, and its time-series signal is used as a node feature to form a graph representation. As a baseline model, a Spatial-Temporal Graph Convolutional Network (ST-GCN), which can simultaneously learn spatial and temporal dependencies in time-series data, is introduced.

Furthermore, in real-world environments where signal responses fluctuate due to noise and where limited training data often lead to unstable estimation of inter-sensor relationships, an extension of ST-GCN incorporating a general spatial attention mechanism, referred to as ST-GCN/SA, is first investigated. However, because self-attention relies on global correlations across the entire dataset, the estimation of sensor relationships does not sufficiently converge under noisy conditions or limited data scenarios, resulting in unstable performance. To overcome these limitations, this study proposes an Anchor-Guided ST-GCN (ST-GCN/AG), which incorporates an anchor-guided mechanism based on node reliability. In the proposed method, node reliability is estimated for each frame, and nodes with high reliability are selected as anchors. Spatial relationships among sensors are then actively guided with respect to these anchors, thereby reducing uncertainty in spatial relationship estimation that cannot be sufficiently suppressed by self-attention alone, and improving training stability and generalization performance.

For evaluation, five representative daily postural actions were selected as target motions: Standing Up, Sitting Down, Sleeping, Getting Up, and Falling. Measurements were conducted under four sensor orientation conditions (east, west, south, and north), with ten repetitions for each condition. The sensor sampling rate was set to 10 Hz, and each trial consisted of a 10 s time series (100 frames), resulting in a total of 1,000 single-action samples. In addition, to evaluate robustness to posture

transitions, paired-action sequences formed by concatenating two actions, as well as long continuous sequences consisting of ten consecutive actions, were constructed. This design enables evaluation not only of isolated actions but also of recognition stability in continuous motion sequences.

In the experiments, the neural network architecture was fixed to a five-layer configuration, and the training data ratio was varied among 50%, 33%, and 25% to analyze the impact of limited data availability on performance. The results show that ST-GCN/SA provides moderate improvements under intermediate data conditions; however, its performance degrades in more limited data scenarios, particularly in single-action classification, indicating insufficient stability. In contrast, the proposed ST-GCN/AG consistently outperforms both the baseline ST-GCN and ST-GCN/SA across all conditions. Notably, under the limited data conditions of 33% and 25%, performance degradation remains relatively small, demonstrating high robustness. These results indicate that the proposed ST-GCN/AG effectively overcomes the limitations of conventional attention mechanisms and exhibits low dependency on training data volume, making it a highly effective approach for sensor-based posture estimation in real-world environments where data acquisition is constrained.

## Abstract

本研究は、屋内生活空間における姿勢推定システムを構築してプライバシーに配慮した非接触センシングにより人の姿勢動作を推定する手法を検討するものである。カメラを用いた姿勢推定は高精度化している一方、常時撮影を伴う運用はプライバシー上の懸念が大きく、住宅や介護環境への実導入には課題が残る。これに対し、24GHz帯の準ミリ波レーダーは暗所や遮蔽に強く、外観情報を取得しないためプライバシー保護の観点で有利である。しかし、準ミリ波レーダーによる観測は空間分解能が低く、反射・ノイズ・環境依存性の影響を受けやすいことから、単一センサの信号のみで姿勢を安定して推定することは難しい。

そこで本研究では、複数台の準ミリ波レーダーを用い、センサ群の時系列観測を「空間（センサ間関係）と時間（動作遷移）」の両面から統合的に扱う学習枠組みを構築した。複数センサをノードとみなし、時系列信号をノード特徴量として与えることでグラフ表現を形成し、空間依存と時間依存を同時に学習可能な Spatial-Temporal Graph Convolutional Network (ST-GCN) をベースラインとして導入した。

さらに、実環境データにおいてノイズ等により信号反応が変動し、少量データ条件下ではセンサ間関係の推定が不安定化しやすいという問題に対し、まず一般的な空間注意機構を導入した ST-GCN/SA による拡張を検討した。しかし、SA はデータ全体の相関に依存するため、ノイズ混入時や少量データ下では関係性の推定が十分に収束せず、性能が振るわないという課題が明らかとなった。そこで、SA の限界を克服する発展的拡張として、ノード信頼度に基づくアンカー誘導機構を組み込んだ Anchor-Guided ST-GCN (ST-GCN/AG) を提案した。提案手法では、各フレームにおけるノードの信頼度を推定し、信頼度の高いノードをアンカーとして選択する。そしてアンカーを起点として他ノードとの関係を能動的に誘導することで、SA では抑制しきれなかった空間関係推定の不確かさを排除し、学習の安定化と汎化性能の向上を狙う。

評価のため、日常生活で代表的な姿勢変化に焦点を当て、Standing Up, Sitting Down, Sleeping, Getting Up, Falling の 5 動作を計測対象とした。センサに対する方位を東西南北の 4 条件に変化させ、各条件 10 回ずつ反復計測を行った。センサのサンプリングレートは 10 Hz とし、1 試行あたり 10 s (100 フレーム) の時系列として単独動作データを構成し、合計 1,000 サンプルを得た。さらに、動作遷移の追従性を評価するために、2 つ動作を連結した組み合わせ動作データおよび 10 組の動作を連続させた長尺連続動作データを構成し、単独動作だけでなく連続系列における識別安定性も検討した。

実験ではニューラルネットワークの構成を 5 層に設定し、学習データ割合を 50%, 33%, 25% に変化させ、学習データ量の制約が性能に与える影響を比較した。その結果、検討段階の ST-GCN/SA は中程度のデータ量では改善が見られるものの、より少量の条件では単独動作の精度が低下し、期待されるほどの安定性は得られなかった。これに対し、上位手法として提案した ST-GCN/AG は、全ての条件でベースラインおよび SA を上回る精度を記録した。特に 33% および 25% の少量データ条件に

において性能低下が小さく, 高い安定性を維持することを確認した. 以上より, 提案する ST-GCN/AG は, 従来の注意機構 (SA) の弱点を克服し, 学習データ量への依存性が小さく, 実環境でデータ取得が制約されるセンサベース姿勢推定において極めて有効な手法であることを示した.

# 目次

<b>第1章</b>	<b>はじめに</b>	<b>1</b>
1.1	本論文の構成	2
<b>第2章</b>	<b>研究の背景と目的</b>	<b>3</b>
2.1	研究背景	3
2.2	研究目的	4
<b>第3章</b>	<b>関連研究</b>	<b>6</b>
3.1	関連研究	6
3.1.1	画像を用いた姿勢推定	6
3.1.2	ミリ波レーダーを用いた姿勢推定	6
3.2	予備知識	7
3.2.1	姿勢推定 (Pose Estimation)	7
3.2.2	準ミリ波レーダー	8
3.2.3	Graph Convolutional Networks (GCN)	8
<b>第4章</b>	<b>準ミリ波レーダーを使用した姿勢推定システム</b>	<b>11</b>
4.1	はじめに	11
4.2	システムの全体構成	12
4.3	センサデータの補正と前処理	12
4.4	機械学習による姿勢推定	13
4.4.1	時空間グラフ畳み込みネットワーク	13
4.4.2	空間アテンションを適用した時空間グラフ畳み込みネットワーク (ST-GCN/SA)	14
4.4.3	アンカー選択に基づく時空間グラフ畳み込みネットワーク (ST-GCN/AG)	16
<b>第5章</b>	<b>姿勢推定システムの構築・実験</b>	<b>19</b>
5.1	実験環境	19
5.1.1	実験機材	19
5.1.2	実験環境	19
5.2	姿勢動作の計測	20
5.2.1	データセットの構成	22

5.3	姿勢推定モデルの構築	23
5.3.1	データ前処理	23
5.3.2	ST-GCN/SAによる動的グラフ学習	24
5.3.3	ST-GCN/AGによる動的グラフ学習	25
5.3.4	学習条件	25
<b>第6章</b>	<b>評価</b>	<b>27</b>
6.1	はじめに	27
6.2	評価手法	27
6.2.1	学習過程の評価	27
6.2.2	単独動作および動作遷移系列に対する評価	27
6.2.3	ネットワーク構成およびデータ分割条件の比較設定	28
6.3	結果	29
6.3.1	データ分割比率による精度の比較	32
6.4	考察	36
6.5	まとめ	37
<b>第7章</b>	<b>おわりに</b>	<b>38</b>
7.1	本研究の概要と成果	38
7.2	今後の課題	39
	謝辞	45
	研究業績	46
	付録	47

# 目次

4.1	システム全体図 . . . . .	12
4.2	ST-GCN/AG の構成 . . . . .	16
5.1	実験環境およびセンサ配置図 . . . . .	20
5.2	姿勢動作計測の様子 (a) 初期化,(b) 動作,(c) 予備,(d) 終了 . . . . .	22
6.1	全体精度: 3 層 . . . . .	30
6.2	全体精度: 5 層 . . . . .	31
6.3	全体精度: 9 層 . . . . .	32
7.1	ST-GCN の学習における損失曲線 . . . . .	47
7.2	ST-GCN/SA の学習における損失曲線 . . . . .	48
7.3	ST-GCN/AG の学習における損失曲線 . . . . .	48
7.4	ST-GCN による単独姿勢動作の混同行列 . . . . .	49
7.5	ST-GCN/SA による単独姿勢動作の混同行列 . . . . .	49
7.6	ST-GCN/AG による単独姿勢動作の混同行列 . . . . .	50
7.7	ST-GCN による組み合わせ姿勢動作 (2 組) の混同行列 . . . . .	50
7.8	ST-GCN/SA による組み合わせ姿勢動作 (2 組) の混同行列 . . . . .	51
7.9	ST-GCN/AG による組み合わせ姿勢動作 (2 組) の混同行列 . . . . .	51
7.10	ST-GCN による組み合わせ姿勢動作 (10 組) の混同行列 . . . . .	52
7.11	ST-GCN/SA による組み合わせ姿勢動作 (10 組) の混同行列 . . . . .	53
7.12	ST-GCN/AG による組み合わせ姿勢動作 (10 組) の混同行列 . . . . .	54
7.13	ST-GCN の学習における損失曲線 . . . . .	55
7.14	ST-GCN/SA の学習における損失曲線 . . . . .	55
7.15	ST-GCN/AG の学習における損失曲線 . . . . .	56
7.16	ST-GCN による単独姿勢動作の混同行列 . . . . .	56
7.17	ST-GCN/SA による単独姿勢動作の混同行列 . . . . .	57
7.18	ST-GCN/AG による単独姿勢動作の混同行列 . . . . .	57
7.19	ST-GCN による組み合わせ姿勢動作 (2 組) の混同行列 . . . . .	58
7.20	ST-GCN/SA による組み合わせ姿勢動作 (2 組) の混同行列 . . . . .	59
7.21	ST-GCN/AG による組み合わせ姿勢動作 (2 組) の混同行列 . . . . .	59
7.22	ST-GCN による組み合わせ姿勢動作 (10 組) の混同行列 . . . . .	60
7.23	ST-GCN/SA による組み合わせ姿勢動作 (10 組) の混同行列 . . . . .	61

7.24	ST-GCN/AGによる組み合わせ姿勢動作（10組）の混同行列 . . . . .	62
7.25	ST-GCNの学習における損失曲線 . . . . .	63
7.26	ST-GCN/SAの学習における損失曲線 . . . . .	63
7.27	ST-GCN/AGの学習における損失曲線 . . . . .	64
7.28	ST-GCNによる単独姿勢動作の混同行列 . . . . .	64
7.29	ST-GCN/SAによる単独姿勢動作の混同行列 . . . . .	65
7.30	ST-GCN/AGによる単独姿勢動作の混同行列 . . . . .	65
7.31	ST-GCNによる組み合わせ姿勢動作（2組）の混同行列 . . . . .	66
7.32	ST-GCN/SAによる組み合わせ姿勢動作（2組）の混同行列 . . . . .	67
7.33	ST-GCN/AGによる組み合わせ姿勢動作（2組）の混同行列 . . . . .	67
7.34	ST-GCNによる組み合わせ姿勢動作（10組）の混同行列 . . . . .	68
7.35	ST-GCN/SAによる組み合わせ姿勢動作（10組）の混同行列 . . . . .	69
7.36	ST-GCN/AGによる組み合わせ姿勢動作（10組）の混同行列 . . . . .	70

# 表 目 次

5.1	収集した姿勢動作データセットの詳細	21
6.1	ネットワーク層数および手法別の姿勢動作認識精度	30
6.2	単独姿勢推定における精度 (学習データ 50%)	33
6.3	組み合わせ姿勢動作 (2 組) における精度 (学習データ 50%)	33
6.4	組み合わせ姿勢動作 (10 組) における精度 (学習データ 50%)	33
6.5	単独姿勢推定における精度 (学習データ 33%)	34
6.6	組み合わせ姿勢動作 (2 組) における精度 (学習データ 33%)	34
6.7	組み合わせ姿勢動作 (10 組) における精度 (学習データ 33%)	34
6.8	単独姿勢推定における精度 (学習データ 25%)	35
6.9	組み合わせ姿勢動作 (2 組) における精度 (学習データ 25%)	35
6.10	組み合わせ姿勢動作 (10 組) における精度 (学習データ 25%)	35
6.11	学習データ分割比率別における全体認識精度	36

# 第1章 はじめに

近年, 屋内空間にセンサを常設し, 継続的に利用者の状態を把握する防犯・監視システムや高齢者見守りシステムにおいて, 人物の姿勢や動作を推定する技術への需要が高まっている. 姿勢推定は, 防犯・監視, 生活支援, リハビリテーション支援など幅広い応用が期待される一方で, 住宅内における常時撮影を前提とするシステムはプライバシー侵害への懸念があり, 課題を抱えている. また, カメラベースの手法は外光等の照明条件や遮蔽物の影響を受けやすく, 安定運用することが難しい場合がある.

このような背景から, 人物の外観情報を取得せずに行動や姿勢を推定可能なセンサベース手法が注目されている. 中でもミリ波レーダーは, 暗所や軽度の遮蔽物下でも動作可能であり, プライバシー保護に優れるという利点を有する. しかし, レーダーから得られる情報は画像と比較して空間分解能が低く, ノイズや多重反射の影響も受けやすい. そのため, 単一センサの信号のみから詳細な姿勢や動作を安定して推定することは容易ではなく, 複数センサの情報統合や時系列情報の活用, ならびに学習モデル設計が重要となる.

本研究では, 24GHz 帯の準ミリ波レーダーを用いた非接触・プライバシー配慮型の姿勢推定システムを対象とし, 複数センサから得られる時系列信号に対して機械学習を適用することで, 日常生活における代表的な姿勢動作を高精度に分類する手法を検討する. 具体的には, 屋内環境におけるセンサ配置とデータ収集 (ハードウェアフェーズ) と, センサデータの補正・前処理および学習モデル構築 (ソフトウェアフェーズ) を一体として設計し, 実環境における推定精度と安定性の観点から評価を行う.

また本研究では, 時系列データの空間・時間依存を同時に扱える時空間グラフ畳み込みネットワーク (ST-GCN) をベースラインとして採用し, センサ間関係をより適応的に学習するための拡張手法を検討する. とりわけ, フレームごとのセンサ信頼度に基づいてアンカーを決定し, そのアンカーを起点としてセンサ間関係を誘導するアンカー選択に基づく時空間グラフ畳み込みネットワーク (ST-GCN/AG) を使用を提案し, 固定グラフに基づく ST-GCN および空間アテンションを適用した ST-GCN/SA との比較により有効性を検証する. さらに, 学習データ量の制約を想定し, 学習データ割合を変化させた条件下での性能変化を分析することで, 実運用を見据えた頑健性についても議論する.

## 1.1 本論文の構成

本論文は以下の 7 章から構成される.

- **第 1 章 (はじめに)** では、本研究の導入として、研究の全体像を概説する.
- **第 2 章 (研究の背景と目的)** では、本研究が取り組む課題の所在を明らかにするため、詳細な研究背景および解決すべき目的を明確にする.
- **第 3 章 (関連研究)** では、既存の姿勢推定技術、準ミリ波レーダーの特性、およびグラフ畳み込みネットワーク (GCN) に関する関連研究を整理し、本研究の基盤となる予備知識について述べる.
- **第 4 章 (準ミリ波レーダーを使用した姿勢推定システム)** では、提案システムの全体構成、センサデータの前処理手法、および提案手法である ST-GCN/SA と ST-GCN/AG の詳細なネットワーク設計について詳述する.
- **第 5 章 (姿勢推定システムの構築・実験)** では、システムの構築に使用した実験環境、姿勢動作データの計測手順とデータセット構成、および具体的な学習条件について述べる.
- **第 6 章 (評価)** では、各種評価実験の結果を提示し、ネットワーク構成やデータ分割条件の違いが推定精度に与える影響について考察を行い、提案手法の有効性を議論する.
- **第 7 章 (おわりに)** では、本研究で得られた成果を総括し、残された課題と今後の展望をまとめる.

## 第2章 研究の背景と目的

本章では、本研究の背景および目的、ならびに関連する先行研究について述べる。はじめに、人物姿勢推定に関する従来研究として、カラー画像を用いた姿勢推定手法と、近年注目されているミリ波レーダーを用いた姿勢推定手法について整理する。次に、ミリ波レーダーの中でも、本研究で対象とする 24GHz 帯準ミリ波レーダーの特性について説明し、高周波帯ミリ波レーダーとの違いを明確にする。その上で、既存研究における課題を整理し、それらの課題に対して本研究のニーズについて説明する。

### 2.1 研究背景

姿勢推定とは、カメラや各種センサを用いて被験者の身体各部位（頭部、胴体、四肢など）の位置関係を推定し、それらをコンピュータ上で扱える座標データとして表現する技術である。得られた骨格情報を時系列的に解析することで、人の行動や状態を認識することが可能となり、防犯・監視システム、高齢者見守り、ゲームなどの幅広い応用が期待されている。

近年、深層学習を中心とした機械学習技術の発展により、カラー画像から人物の骨格を推定して姿勢を高精度に推定する技術が急速に発展している [1],[2]。

一方で、画像を用いた姿勢推定は、画像から人物の特徴を直接取得するため、プライバシー侵害の懸念が大きい。特に住居等のプライベートな空間においては、常時カメラによる撮影を行うことがプライバシーの懸念等から受け入れられにくいという課題がある。

また、推定にカメラを使用するためカメラ機器の特性上照明や遮蔽物などの撮影環境に強く影響される。

このような背景から、近年ではミリ波レーダーをはじめとするセンサベースの姿勢推定手法が注目を集めている。ミリ波レーダーは、電波の反射特性を利用して対象物の位置や動きを検出するため、人物の外観や顔画像を取得することなく行動や姿勢の情報を抽出できる。この特性により、プライバシーを保護しつつ人の動作を認識できる点が大きな利点である。さらに、暗所や煙、軽度の遮蔽物が存在する環境でも動作可能であることから、カメラベース手法では困難な状況下での利用も期待されている。

しかしながら、ミリ波レーダーによって得られる情報は、画像と比較して空間分

解能が低く、ノイズや多重反射の影響を受けやすい。そのため、レーダー信号から直接人体の詳細な骨格構造を推定することは容易ではなく、姿勢推定精度や汎用性の向上が依然として課題である。これに対し、機械学習、特に時空間的な構造を扱えるモデルを用いることで、低解像度なセンサデータから人の姿勢や行動を推定する研究が活発に行われている [3],[4],[5]。

本研究では、このような背景を踏まえ、プライバシーに配慮した姿勢推定技術の実現を目的として、ミリ波レーダーから得られるデータを用いた人物姿勢推定手法について検討する。特に、センサ配置や時系列情報を考慮した機械学習モデルを構築し、従来の画像ベース手法に代替し得る新たな姿勢推定アプローチの可能性を明らかにすることを目指す。

## 2.2 研究目的

本研究の目的は、ミリ波レーダーセンサを用いた非接触型の人物姿勢推定システムを構築し、プライバシーに配慮した室内向け行動分析手法を確立することである。近年、家庭や医療・介護施設などの住宅空間において、人の行動や状態を把握する技術への需要が高まっている。特に、高齢者の見守りや生活支援、リハビリテーション支援といったヘルスケア分野では、被験者に負担を与えず、かつ継続的に利用可能な計測手法が求められている。

従来、人物の姿勢推定にはカメラを用いた画像ベースの手法が広く用いられてきたが、住宅空間における常時撮影はプライバシー侵害の懸念が大きく、社会的受容性の面で課題がある。また、照明条件や遮蔽の影響を受けやすく、実環境での安定運用が難しいという問題も存在する。これらの背景から、人物の外観情報を取得せずに動作や姿勢を推定可能なセンサベースの姿勢推定に注目が集まっている。

センサベースの姿勢推定は近年、低解像度の赤外線画像 [6] や人感センサ、Wi-Fi のチャンネル状態情報 (CSI: Channel State Information) [7] といった多様なデータソースが活用されている。これらの手法では、機械学習処理を導入してセンサデータ特有の特性や環境ノイズを適応的に学習させることで、カメラを用いない非接触かつ高精度な姿勢推定を実現している。

その中でミリ波レーダーを使用したの研究が活発である [8],[9],[10]。ミリ波レーダーは、電波の反射特性を利用して対象物の距離や動きを検出するセンサであり、暗所や遮蔽物が存在する環境でも動作可能であるという利点を持つ。一方で、取得されるデータは画像と比較して空間分解能が低く、人体構造を直接的に捉えることは困難である。そのため、住宅等の室内空間において姿勢推定を行うためには、ミリ波レーダー固有の特性を十分に理解し、センサ配置や信号データの処理、推定アルゴリズムを総合的に設計する必要がある。

本研究ではまず、住宅空間を想定した環境下において、ミリ波レーダーによる人物検出および姿勢推定に影響を与える要因を明らかにする。具体的には、センサの設置高さや設置角度、壁面・天井への配置の違い、ならびに使用するセンサの個数

が、取得される信号や推定精度に与える影響を比較・検討する。これにより、室内環境において安定した姿勢推定を行うための最適なセンサ配置条件を導出することを目的とする。

次に、ミリ波レーダーから得られる低解像度かつノイズを含むデータに対し、機械学習を用いた補正および推定手法を適用する。特に、時系列情報や複数センサ間の関係性を考慮することで、単一フレームでは困難な姿勢推定精度の向上を図る。これにより、レーダーデータのみから人物の姿勢や動作状態を高精度に推定可能なソフトウェア処理手法の確立を目指す。

さらに、本研究では推定された姿勢情報を基に、ユーザーの行動分析を行う。姿勢の変化や遷移を解析することで、立位・座位・臥位といった基本姿勢や日常動作を識別し、生活行動の把握につなげる。このような行動分析結果を活用することで、カメラを使用せずに利用者の状態を把握可能な、プライバシーに配慮したシステムの提案を行う。

本研究では、このような 24GHz 帯準ミリ波レーダーの特性を踏まえ、単一センサの性能に依存するのではなく、複数センサの配置や時系列情報の活用、ならびに機械学習による補完を組み合わせることで、人物姿勢推定の精度向上を図る。高性能なハードウェアに依存せず、汎用的かつ導入容易なセンサを用いて実用的な姿勢推定を実現する。ミリ波レーダーを用いた非接触型姿勢推定の基盤技術を確立するとともに、実環境での利用を想定したセンサ配置設計および機械学習による推定精度向上手法を提示することを目的とする。

# 第3章 関連研究

## 3.1 関連研究

### 3.1.1 画像を用いた姿勢推定

画像を用いた姿勢推定は、RGB 画像や動画を入力として、人体の関節位置（キーポイント）を推定する技術であり、近年活発に研究されている。

代表的な手法として、OpenPose が挙げられる [1]。OpenPose は、畳み込みニューラルネットワーク（CNN）を用いて、人体の各関節位置とそれらの接続関係を同時に推定する手法であり、複数人物に対してもリアルタイムで高精度な姿勢推定を実現している。また、画像全体から関節候補を先に抽出し、それらを統合して個々の人物の姿勢を再構成するボトムアップ型の手法を採用しているため、人物数に依存しない効率的な処理が可能である。

OpenPose では、各関節位置の推定において、画像から関節信頼度マップと呼ばれるヒートマップを生成する。この信頼度マップは、各画素において特定の関節（例：肩や膝）が存在する確率を表しており、そのピーク位置が関節候補として抽出される。さらに、関節同士の対応関係を推定するために、Part Affinity Fields（PAFs）と呼ばれるベクトル場が同時に推定される。PAFs は、各関節ペア（例：肩と肘）を結ぶ方向と強さを表す 2 次元ベクトル場であり、これにより複数人物が複雑に重なり合う場合でも、同一人物に属する関節同士を適切に接続することが可能となる。これらの信頼度マップと PAFs を統合することで、人体の骨格構造（スケルトン）を高精度に再構成することができる。

しかしながら、画像ベースの手法には実環境においていくつかの課題が存在する。まず照明条件や遮蔽の影響である。暗所や逆光環境下では画質が著しく低下し、検出の精度が低下する。また、屋内環境においては家具や壁による遮蔽が生じやすく、安定した姿勢推定が困難となる場合がある。そして、プライバシー保護の問題が挙げられる。カメラによる撮影は、個人の顔や服装、室内の詳細な状況を不可避免的に記録するため、プライバシーの問題は避けられない。

### 3.1.2 ミリ波レーダーを用いた姿勢推定

ミリ波レーダーを用いた研究の多くは、レーダーから得られる高精度な距離情報、ドップラー情報、あるいはそれらを可視化した時間-距離（Time-Range）マップや

時間-ドップラー (Time-Doppler) マップを特徴表現として用い、機械学習や深層学習により行動分類を行っている [10],[11],[9]. 特に、畳み込みニューラルネットワーク (CNN) を用いてレーダーデータを画像的に処理する手法が多く提案されており、歩行や運動動作などの限られた行動クラスに対して高い認識精度が報告されている.

これらの手法の多くは、高精度な 60GHz 帯以上のミリ波レーダーを使用しており、高い距離分解能や角度分解能を前提としたシステム構成となっている. このような高周波帯ミリ波レーダーは、人体の微細な動きや形状を捉える上で有効である. しかし、センサのコストや周波数帯に関する法規制や地域差への配慮が必要となる場合もあり、住宅環境への常設や広範な普及を想定した場合には導入のハードルがある.

さらに、ミリ波レーダーから得られるデータは、画像と比較して空間分解能が低く、ノイズや反射の影響を受けやすいという特性がある. そのため、単一センサの情報のみを用いて高精度な姿勢推定を行うことは難しく、複数センサの情報統合や時系列情報の活用が重要な課題となっている. 特に、人体構造やセンサ間の関係性をどのようにモデル化し、推定精度の向上につなげるかが、ミリ波レーダーを用いた姿勢推定における重要な研究課題である.

## 3.2 予備知識

### 3.2.1 姿勢推定 (Pose Estimation)

姿勢推定とは、画像や動画、あるいは各種センサデータを用いて、人物の身体各部位の位置関係や姿勢を推定する技術である. 一般に、人体は頭部、胴体、上肢および下肢といった複数の部位から構成されており、これらの端点や関節位置は「骨格点 (キーポイント)」として表現される. 姿勢推定は、これらのキーポイントをコンピュータ上の座標として推定することで、人体の構造や動作を定量的に扱うことを可能にする.

特にコンピュータビジョン分野においては、カラー画像や動画を入力とし、人物の骨格構造を推定する画像ベースの姿勢推定技術が盛んに研究されている. 従来の姿勢推定手法では、背景差分法やエッジ情報、形状モデルを用いたテンプレートマッチングなどが主流であり、事前に定義された人体モデルとの対応付けによって姿勢を推定していた. しかし、これらの手法は、照明条件の変化や背景の複雑化、人物同士の重なりなどに弱く、実環境での適用には限界があった.

近年では、深層学習 (Deep Learning) の発展により、姿勢推定の精度は飛躍的に向上している. 畳み込みニューラルネットワーク (CNN) を基盤とした手法により、画像中の人物領域から各キーポイントの位置を直接回帰することが可能となり、複雑な背景や遮蔽 (オクルージョン) が存在する環境下においても、頑健な推定が実現されている. これらの手法は、単一人物を対象とした姿勢推定にとどまらず、複数

人物が同時に存在するシーンにおいても適用可能であり, 行動認識や人流解析などへの応用が進められている.

一方で, 画像ベースの姿勢推定には, 人物の外観情報を直接取得する必要があるという特性上, プライバシー侵害の懸念が伴う. 特に住宅内や医療・介護施設などの私的空間においては, 常時カメラによる撮影を行うことが社会的に受け入れられにくく, 実運用における大きな課題となっている. また, 照明条件やカメラ視点に強く依存する点も, 環境適応性の観点から問題点として挙げられる.

このような課題を背景として, 画像を用いないセンサベースの姿勢推定手法が注目されている. センサベース手法では, LiDAR やミリ波レーダーなどを用いて人物の位置や動きを検出し, その情報から姿勢や行動を推定する. LiDAR は高い距離分解能を有する一方で, 装置の大型化やコストの問題があり, 住宅環境への導入には制約がある.

これに対し, ミリ波レーダーは電波の反射特性を利用することで, 暗所や遮蔽物が存在する環境下でも動作可能であり, 人物の外観情報を取得しないという点でプライバシー保護に優れている. しかし, ミリ波レーダーから得られる情報は画像と比較して空間分解能が低く, 人体の詳細な骨格構造を直接観測することは困難である. そのため, センサデータをどのように解釈し, 姿勢推定へと結び付けるかが重要な研究課題となっている.

### 3.2.2 準ミリ波レーダー

本研究では, 人物の姿勢推定を行うためのセンサとして, 24 GHz 帯の準ミリ波レーダーを用いる. 準ミリ波レーダーとは, 一般に 24 GHz 帯前後の周波数を用いる電波センサを指し, 60 GHz 帯以上の高周波ミリ波レーダーと比較して波長が長く, 空間分解能は低いものの, 低コストかつ低消費電力で実装可能であるという特徴を有する.

60GHz 帯以上の高周波ミリ波レーダーでは, 使用条件や出力制限, 地域ごとの規制に配慮する必要がある場合が多く, 研究段階から実運用への移行において障壁となることがある. これに対し, 24GHz 帯は ISM(Industrial, Scientific and Medical) バンドとして国際的に割り当てられている周波数帯であり, 法規制への対応が容易である.

### 3.2.3 Graph Convolutional Networks (GCN)

グラフ構造データに対する畳み込み演算は, 近年のグラフ深層学習の発展に伴い多くの手法が提案されている. 初期の研究としては, グラフ上で畳み込みを実現するために, 分子構造を対象とした畳み込みモデルを提案 [12] や, グラフの局所構造に基づく畳み込みを定義した手法 [13] が挙げられる.

その後、グラフラプリアンに基づくスペクトル畳み込みを近似することで、計算効率と表現能力を両立した Graph Convolutional Network (GCN) を提案した [14]. このモデルでは、隣接ノードの特徴量を正規化して集約することで、ノード表現を更新する枠組みが確立された。

さらに、グラフ畳み込みの枠組みは空間的観点からも発展しており、拡散過程に基づく Diffusion Convolutional Neural Networks[15] や、大規模グラフに対して帰納的に表現学習を行う GraphSAGE[16] などが提案されている。これらの手法は、ノードの近傍情報を直接集約することで、スケーラビリティと汎用性を向上させている。

また、グラフ構造データは非ユークリッド空間に属するため、幾何学的深層学習の観点からの拡張も行われており、多様体上での畳み込みを可能とするフレームワークも提案されている [17].

さらに、グラフ構造そのものを推定対象とする研究として、ノード間の相互作用を潜在変数として学習する Neural Relational Inference[18] を提案しており、グラフ構造の動的モデリングへと発展している。

このように、GCN は初期のグラフ畳み込みの定式化から、空間的手法、幾何学的拡張、さらには関係推論へと発展しており、関係性を本質とするデータに対する有効な表現学習手法として広く利用されている。

一般に、グラフはノード（頂点）とエッジ（辺）から構成される。GCN では、各ノードが特徴ベクトルを持ち、エッジはノード間の関係性を表現する。学習過程においては、あるノードの特徴量が隣接ノードの特徴量と集約・更新されることで、より高次元な表現が獲得される。この操作はグラフ畳み込みと呼ばれ、ノード間の接続関係を反映した特徴抽出を可能にする。

代表的な GCN の層は、隣接行列を用いて次式のように表される。

$$H^{(l+1)} = \sigma \left( \tilde{D}^{-\frac{1}{2}} \tilde{A} \tilde{D}^{-\frac{1}{2}} H^{(l)} W^{(l)} \right) \quad (3.1)$$

ここで、 $H^{(l)}$  は  $l$  層目のノード特徴行列、 $W^{(l)}$  は学習可能な重み行列、 $\tilde{A}$  は自己ループを加えた隣接行列、 $\tilde{D}$  はその次数行列、 $\sigma$  は活性化関数を表す。この式は、隣接ノードの特徴量を正規化した上で集約する処理を意味している。

GCN の基本的な原理は、各ノードの特徴をその近傍ノードからの情報を集約することにより、より表現力の高い特徴へと更新していく点にある。このため、ノードの局所的な構造だけでなく、層を重ねることでグラフ全体の構造情報を間接的に学習することが可能となる。この性質により、GCN はソーシャルネットワーク解析や分子構造解析など、関係性が本質的な役割を果たすタスクに広く用いられている。

姿勢推定や行動認識の分野においては、人体構造をグラフとして表現できる点が GCN の大きな利点である。人体の関節や部位をノードとし、骨格の接続関係をエッジとして定義することで、人の姿勢はグラフ構造データとして扱うことが可能となる。この表現により、関節間の物理的な関係性を保ったまま特徴量を学習できるた

め, 人体の構造的制約を考慮した姿勢推定が可能となる.

さらに,GCN は時系列データにも適用可能であり, 時間方向の情報を組み合わせることで, 動作や行動といった連続的な現象を扱うことができる. この考え方に基づき, 空間的な骨格構造と時間的な変化を同時に扱う手法として, Spatial-Temporal Graph Convolutional Network (ST-GCN) が提案されている [19]. ST-GCN では, 空間方向のグラフ畳み込みと時間方向の畳み込みを組み合わせることで, 姿勢構造と動作の時間変化を同時に学習することが可能であり, 姿勢推定や行動認識において高い性能を示している.

# 第4章 準ミリ波レーダーを使用した姿勢推定システム

## 4.1 はじめに

本研究では, 準ミリ波レーダーを用いた非接触・プライバシー配慮型の姿勢推定システムを実現するための手法を提案する. 従来の姿勢推定研究の多くはカメラ画像を入力とする手法に基づいており, 高精度な骨格推定が可能である一方, 撮影によるプライバシー侵害の懸念や, 照明条件・遮蔽物の影響を受けやすいという課題を有している. これに対し, ミリ波レーダーは被写体の外観情報を取得せず, 環境変化にも比較的頑健であることから, 屋内環境における人物状態把握のための有効なセンシング手段として注目されている.

しかし, 準ミリ波レーダーは距離分解能や空間分解能が低く, 単純なルールベース処理や単一センサ構成では, 人物の詳細な姿勢や動作を推定することが困難である. また, センサ設置位置や屋内反射の影響により, 取得される信号にはばらつきやノイズが含まれるため, センサ特性を考慮した補正および学習手法が不可欠である.

そこで本研究では, 提案手法をハードウェアフェーズとソフトウェアフェーズの二段階に分けて構成する. まずハードウェアフェーズでは, 実際の屋内環境に準ミリ波レーダーを設置し, 人物の動作時に得られるセンサ出力を計測することで, 距離ゾーンごとの反射強度分布や設置位置による特性の違いを明らかにする. これにより, 居住空間における適切なセンサ個数および配置条件を検討する.

次にソフトウェアフェーズでは, ハードウェアフェーズで得られたセンサ特性を基に, センサデータの補正および前処理を行う. さらに, センサデータ取得と同時にカメラによる撮影を行い, 取得した画像に対して既存の骨格推定手法を適用することで, 姿勢推定モデルの学習に用いる教師データを生成する. これらのセンサ情報と教師データを用い, グラフ畳み込みネットワーク (GCN) を適用することで, 時系列的かつ空間的なセンサ情報の関係性を考慮した姿勢推定を実現する.

本章では, 以上の方針に基づく提案手法の全体構成を示し, 続く節においてセンサ構成, データ処理手法, および学習モデルの詳細について順に述べる.

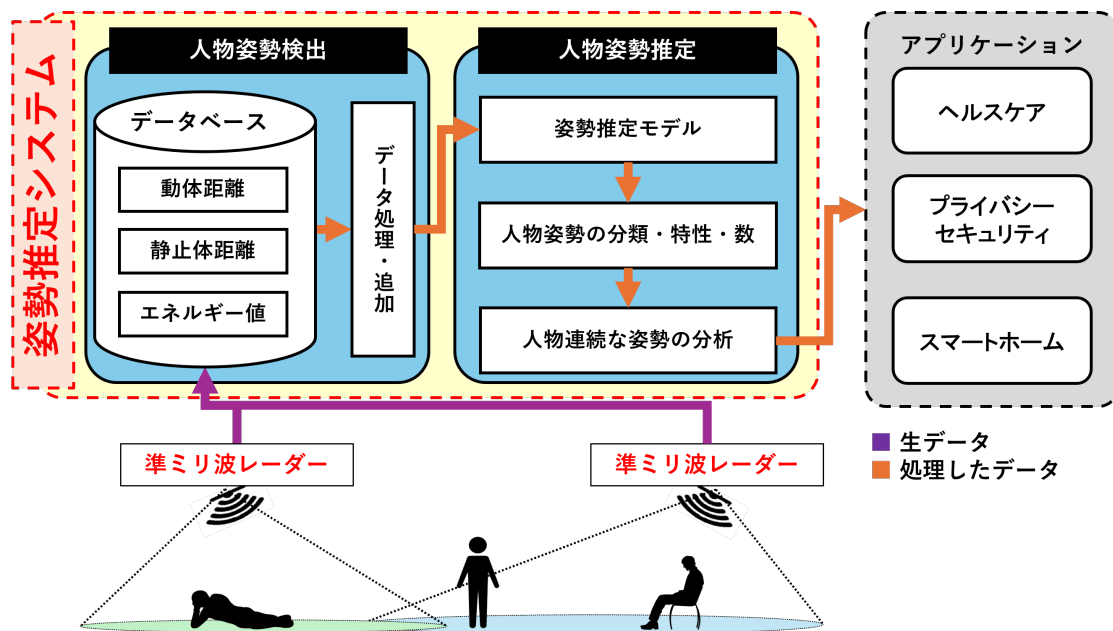


図 4.1: システム全体図

## 4.2 システムの全体構成

以上のセンサ構成および配置により, 本研究では屋内環境における準ミリ波レーダーの特性を考慮しつつ, 姿勢推定に適したセンサデータの収集を行う. 次節では, これらのセンサデータに対する補正および前処理手法について述べる.

## 4.3 センサデータの補正と前処理

準ミリ波レーダーセンサから取得されるデータは, 人物の姿勢や動作に関する有用な情報を含む一方で, ノイズや信号の揺らぎ, センサ配置の違いによる感度差などの影響を受けやすいという課題がある. これらは姿勢推定精度の低下要因となるため, 本研究では学習処理を行う前段階として, レーダーセンサデータに対する補正および前処理を行う.

準ミリ波レーダーセンサは, 距離方向に分割された複数のゾーンに対して, 反射強度や距離情報を時系列データとして出力する. これらのデータはゾーンごとに独立して取得されるため, 人物の姿勢変化はゾーン間のエネルギー分布の変化として現れる. 本研究では, これらの距離ゾーン情報を空間的特徴量として扱い, 時系列方向に連結することで, 姿勢変化を表現する特徴系列を構成する.

まず, センサ出力データに対して欠損値および外れ値の補正を行う. 通信遅延や一時的な検出失敗により一部のゾーン情報が欠落した場合には, 前後フレームの値を用いた補間処理を適用する. また, 反射エネルギー値が突発的に大きく変動する

場合には、移動平均処理を用いることで短時間ノイズを抑制する。これにより、人物の動作に起因する変化と誤反応に由来するノイズ成分を分離しやすくする。

次に、センサ配置や個体差による出力値のばらつきを抑えるため、距離ゾーンごとに正規化処理を行う。反射エネルギー値は一定範囲で出力されるが、人物とセンサとの相対位置や姿勢によって分布が異なるため、各ゾーンの値を統一的なスケールに変換する。この正規化により、異なる計測条件下で取得されたデータを同一モデルで扱うことが可能となる。

さらに、本研究ではセンサデータを姿勢推定モデルに入力するため、時系列データを一定フレーム長に分割し、固定長の入力系列として整形する。これにより、人物動作の開始・終了タイミングが異なる場合でも、モデルが時系列パターンとして学習できるようになる。整形後のデータは、距離ゾーンをノード、反射エネルギーや距離情報を特徴量とするグラフ構造として表現され、次節で述べるグラフ畳み込みネットワークへの入力として使用される。

以上の補正および前処理により、準ミリ波レーダーセンサ特有の誤反応ノイズや条件依存性を低減し、姿勢推定に適したセンサ特徴量を生成する。次節では、これらの前処理済みデータを用いた姿勢推定モデルおよび学習手法について詳述する。

## 4.4 機械学習による姿勢推定

### 4.4.1 時空間グラフ畳み込みネットワーク

時系列データの時間的依存関係をモデル化する手法として、時間畳み込みネットワーク (Temporal Convolutional Network: TCN) が提案されている [20]。TCN は、時間軸に沿った次元畳み込みを用いて系列データを処理する手法であり、再帰型ニューラルネットワーク (RNN) とは異なり並列計算が可能である点、および長期依存関係を安定して学習できる点に特徴を持つ。

TCN では、各時刻の出力は有限長の受容野に基づいて計算されるが、畳み込み層を多段に積み重ねることで広い時間範囲の情報を取り込むことが可能となる。これにより、動作の開始から終了に至るまでの時間的な変化パターンを効率的に捉えることができる。

一方で、TCN は時間方向の依存関係を捉えることには優れているが、入力データ内の空間的な構造 (例えば人体関節間の関係やセンサ間の配置) を直接扱うことはできない。

このような空間構造を扱うための手法として、グラフ畳み込みネットワーク (Graph Convolutional Network: GCN) が提案されている。グラフ上の畳み込み演算は、スペクトルグラフ理論に基づき CNN を非ユークリッド構造へ拡張した研究 [21] に端を発する。

これらの時間モデリング (TCN) と空間モデリング (GCN) を統合したモデルが、時空間グラフ畳み込みネットワーク (ST-GCN) である。ST-GCN は、時系列

データにおける空間的構造と時間的遷移を同時に学習するために提案された手法であり、骨格情報に基づく行動認識を主な対象として開発された。

ST-GCNでは、人体骨格を関節点（ノード）と、それらの物理的接続関係（エッジ）からなるグラフとして表現し、各時刻の骨格情報を時間方向に連結することで時空間グラフを構成する。このグラフに対して、空間方向のグラフ畳み込み（Spatial GCN）と時間方向の畳み込み（Temporal Convolution）を組み合わせて適用することで、関節間の相互作用と動作の時間的変化を同時にモデル化する。

ST-GCNにおける空間グラフ畳み込みは、あらかじめ定義された隣接行列に基づいて行われる。グラフ畳み込みでは、ノード近傍の特徴を集約することで局所的な構造を抽出するが、原著では人体骨格の物理構造に基づく隣接行列を用いるとともに、ノード間の距離関係に応じて複数のサブグラフへ分割する手法が採用されており、各関節の役割を考慮した特徴抽出が可能となっている。

一方で、ST-GCNではノード間の接続関係が時間を通じて不変であると仮定されており、隣接行列は固定的に与えられる。この設計は人体骨格のように構造が安定している対象に対しては有効であるが、入力信号の相関構造が時間的に変化する場合には、表現能力に制約が生じる可能性がある。

特に、準ミリ波レーダーから得られる信号は情報量が限定的であり、本研究では複数のセンサを組み合わせることで姿勢推定を行う必要がある。その際、センサの配置によっては特定の姿勢においてレーダー反応が弱くなる場合や、逆に特徴的な反応を示す場合が存在する。このような計測条件下では、センサ間の相関構造そのものが姿勢や動作に応じて時間的に変動するという特徴を持つ。

そのため、固定的な隣接行列に基づく従来のST-GCNでは、動作ごとに重要となる空間的関係の変化を十分に表現できない場合がある。

そこで本研究では、ST-GCNの枠組みを準ミリ波レーダーを用いた姿勢推定に適用しつつ、空間構造の扱いを拡張する。具体的には、各ノードを人体関節ではなく「センサ単位」として定義し、距離ゾーン方向の反射エネルギー分布や距離情報を特徴量として与える。さらに、物理的なセンサ配置に基づく構造をアンカーとして導入し、データ駆動型のアテンション機構によってセンサ間関係を適応的に学習するアンカー選択に基づく時空間グラフ畳み込みネットワーク（ST-GCN/AG）を提案する。

#### 4.4.2 空間アテンションを適用した時空間グラフ畳み込みネットワーク（ST-GCN/SA）

従来のST-GCNでは、ノード間の接続関係はあらかじめ定義された静的隣接行列により固定されているが、近年ではノード間の関係性をデータから動的に学習する手法として、自己注意機構（Self-Attention）を用いたグラフ表現学習が提案されている。代表的なものとして、Graph Attention Network（GAT）[22]があり、ま

た自己注意機構そのものは Transformer において提案された Attention Is All You Need[23] に基づくものである。これらの手法ではノード間の相関をデータ駆動的に推定することで、柔軟な関係モデリングを実現している。

本研究では、これらのアテンションベース手法に着想を得て、ST-GCN に空間アテンション機構を導入した ST-GCN/SA を実装した。本手法では、各サンプルの特徴量から全ノード間の動的な相関関係を直接推定することで、固定グラフに依存しない空間表現を構築する。

まず、入力特徴量  $x \in \mathbb{R}^{C \times T \times V}$  に対し、時間方向に平均化を行うことで、ノード間の空間的関係性に着目した特徴マップ  $x_s$  を生成する。次に、 $1 \times 1$  畳み込み層を用いて Query ( $Q$ ) および Key ( $K$ ) を生成し、それらの内積に基づくドット積アテンションにより動的関係行列  $A_{\text{attn}}$  を算出する：

$$Q = W_q x_s, \quad K = W_k x_s, \quad (4.1)$$

$$A_{\text{attn}} = \gamma \cdot \text{Softmax} \left( \frac{Q^\top K}{\sqrt{d}} \right). \quad (4.2)$$

ここで  $\gamma$  は学習可能なスカラーパラメータであり、初期値  $\gamma = 0$  とすることで、学習初期には静的グラフに基づく安定した特徴抽出を行い、学習の進行に応じて動的関係の寄与を段階的に導入する設計としている。

従来の固定グラフ構造とは異なり、本手法 (ST-GCN/SA) では自己注意機構 (Self-Attention) を用いることで、各サンプルの特徴量から全ノード間の動的な相関関係を直接抽出する。まず、入力特徴量  $x \in \mathbb{R}^{C \times T \times V}$  に対し時間方向に平均化を行い、空間特徴マップ  $x_s$  を生成する。これに対し  $1 \times 1$  畳み込み層を用いて Query ( $Q$ ) および Key ( $K$ ) への写像を行い、ドット積アテンションを適用することで動的関係行列  $A_{\text{attn}}$  を得る：

$$Q = W_q x_s, \quad K = W_k x_s, \quad (4.3)$$

$$A_{\text{attn}} = \gamma \cdot \text{Softmax} \left( \frac{Q^\top K}{\sqrt{d}} \right). \quad (4.4)$$

ここで  $\gamma$  は学習可能なスカラーパラメータであり、初期値 ( $\gamma = 0$ ) から学習を通じて動的情報の寄与度を自動的に調整する。

### 動的グラフ畳み込みと特徴統合

ST-GCN ブロック内では、事前に定義された静的隣接行列  $A_{\text{static}}$  と、算出された動的行列  $A_{\text{attn}}$  を統合することで、物理的な接続関係と状況依存の相関関係の両方を考慮したグラフ構造  $A_{\text{total}}$  を構築する。

$$A_{\text{total}} = A_{\text{static}} + A_{\text{attn}}. \quad (4.5)$$

この統合された隣接行列を用いて空間方向のグラフ畳み込み（GCN）を行い、その後、時間的畳み込み（TCN）および残差接続（Residual Connection）を適用することで、最終的な特徴出力  $f_{out}$  を得る：

$$f_{conv} = (A_{total} \otimes W_{gcn})f_{in}, \quad (4.6)$$

$$f_{out} = \text{ReLU}(\text{TCN}(f_{conv}) + f_{res}). \quad (4.7)$$

これにより、特定の動作において重要となるセンサ（ノード）間の関係性を、データ駆動型で適応的に学習することが可能となる。

#### 4.4.3 アンカー選択に基づく時空間グラフ畳み込みネットワーク（ST-GCN/AG）

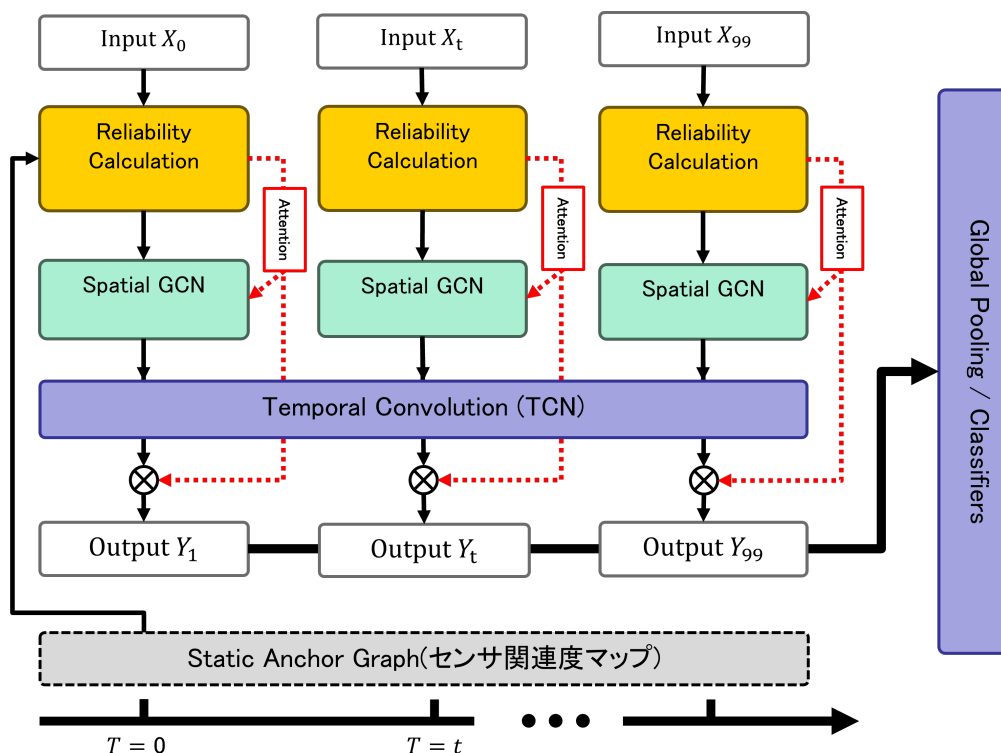


図 4.2: ST-GCN/AG の構成

本研究で提案するアンカー選択に基づく時空間グラフ畳み込みネットワーク（ST-GCN/AG）は、全結合的な静的グラフ構造と、ノードの信頼度に基づいて動的に計算されるアテンション機構を統合したアーキテクチャである。

近年では、ノード間の関係性をデータから動的に学習する手法として、Adaptive Graph Convolutional Network（AGCN）に基づく手法が提案されている。代表的

なものとして、Two-Stream Adaptive Graph Convolutional Network (2s-AGCN) [24] があり、固定グラフに加えて学習可能な動的グラフを導入することで、ノード間の関係をデータ駆動的に最適化し、より柔軟な関係モデリングを実現している。

一方で、これらの手法では、ノード間の関係は主に特徴の類似度に基づいて学習されるため、ノイズを含む入力に対して関係推定が不安定となる場合がある。

これに対し、本研究では、センサ信頼度という物理的関連性に基づいてアンカーを決定し、そのアンカーを起点としてノード間の関係性を構築することで、関係推定に明示的な制約を導入する点に特徴がある。

## 静的グラフの定義

まず、ノード間の基礎的な接続関係を規定する静的隣接行列  $A_{\text{static}}$  を定義する。本実装ではノード数  $V = 3$  の準ミリ波レーダーセンサに対し、自己ループ  $A_0 = I$  と、他ノードへの全結合接続  $A_1 = 1 - I$  を定義し、正規化を行ったものを積み重ねて使用する。これにより、センサ間の基本的な情報共有パスを保証する。

## 各ノード信頼度に基づくアンカー決定

動作の局面に応じて有効な情報を持つセンサ（アンカー）を特定するため、入力データから各ノードの信頼度を推定する。入力特徴  $x$  のチャンネル平均であるエネルギー系列  $e$  に対し、時間窓  $w$  ( $w = 9$ ) における移動分散  $\text{Var}_t(e)$  を計算する。これに対し、温度パラメータ  $\tau$  を用いた Softmax を適用することで、信頼度  $S \in \mathbb{R}^{N \times T \times V}$  を算出する：

$$S_{n,t,v} = \text{softmax}\left(\frac{\text{Var}_t(e_{n,:,v})}{\max(\tau, 10^{-6})}\right). \quad (4.8)$$

この指標により、時間的な信号変動が大きく、動作特徴が顕著なセンサが高い信頼度（アンカー候補）として扱われる。

## アンカー誘導型アテンション (Anchor-Guided Attention)

従来の Dot-product Attention とは異なり、本手法では信頼度  $S$  を用いて動的に「アンカー特徴」を生成し、それとの類似度に基づいて関係性を構築する。まず、時間方向に平均化した信頼度  $\bar{S}$  を重みとして、チャンネル平均特徴  $x_s$  の加重平均をとることで、各サンプルの代表特徴ベクトル（アンカー特徴）  $f_{\text{anchor}}$  を生成する：

$$f_{\text{anchor}} = \sum_v \bar{S}_v \cdot x_{s,v}. \quad (4.9)$$

次に、このアンカー特徴と各ノードの特徴  $x_s$  とのコサイン類似度  $\text{sim}$  を計算し、ReLU 関数と信頼度  $S$  による重み付けを経て動的隣接行列  $A_{\text{anchor}}$  を得る：

$$A_{\text{anchor}} = \lambda \cdot \text{norm} \left( \text{ReLU}(\bar{S} \cdot \text{sim}(f_{\text{anchor}}, x_s)) \right). \quad (4.10)$$

ここで  $\lambda$  はアンカーの影響度を制御する係数 ( $\lambda = 0.8$ ) である。

## ハイブリッドストリームとゲーティング

ST-GCN ブロック内では、静的グラフと動的グラフによる畳み込みを並列に行い、統合する。

$$f_{\text{conv}} = \sum_k (A_{\text{static}}^{(k)} M^{(k)}) f_{\text{in}} W_{\text{static}}^{(k)} + A_{\text{anchor}} f_{\text{in}} W_{\text{dynamic}}. \quad (4.11)$$

ここで  $M$  はエッジ重要度の学習可能マスクである。さらに、算出された特徴量に対し、時間的畳み込み (TCN) を適用した後、信頼度  $S$  の最大値に基づいたゲーティング処理を行うことで、ノイズの多いフレームやセンサの影響を最終出力において抑制する：

$$f_{\text{out}} = \text{ReLU}(\text{TCN}(f_{\text{conv}}) \odot \mathcal{G}(S) + f_{\text{res}}). \quad (4.12)$$

$\mathcal{G}(S)$  は信頼度マップを特徴マップサイズにリサイズしたゲート係数である。

# 第5章 姿勢推定システムの構築・実験

## 5.1 実験環境

### 5.1.1 実験機材

本研究では、人物の姿勢推定を非接触かつプライバシーに配慮した形で実現するため、24GHz 帯準ミリ波レーダーである HLK-LD2410B (Hi-Link 社製) を採用した。本センサは単価が約 2 ドルと非常に安価かつ小型 (35mm × 7mm) であり、UART 通信を介してコンピュータと容易に接続可能であることから、居住空間や公共空間への常設および複数台を用いたシステム構築に適している [25]。

HLK-LD2410B は、最大約 6m の検出距離を持ち、人の存在検知および動体・静止体の識別が可能である。本研究では、屋内における人物の姿勢変化を詳細に捉えるため、センサ前方の検出範囲を最大 9 つの距離ゲート (Z0~Z8) に分割し、各ゾーンにおける反射エネルギー値 (0~100) および距離情報を時系列データとして取得する。この距離ゾーン構造を利用することで、人物の位置変化や姿勢変化に伴う反射分布の推移を捉えることが可能となる。

カメラや LiDAR と比較すると、本センサは空間分解能が低く、単一センサのみでは人体の詳細な形状や骨格構造を直接的に推定することは困難である。また、60GHz 帯や 77GHz 帯の高周波ミリ波レーダーは高い距離・角度分解能を有する反面、高コスト化や消費電力の増大、法規制の課題が存在する。これに対し、24GHz 帯準ミリ波レーダーは分解能の制約はあるものの、プライバシー保護とコストの観点から実用性が高く、複数台の連携によって空間情報の粗さを補完するアプローチが有効であると考えられる。

### 5.1.2 実験環境

居住空間を想定した屋内環境においてデータ収集実験を行った。実験場所は学内の廊下を使用し、幅 1.93m、奥行 4.0m、高さ 2.4m の空間を計測領域とした。準ミリ波レーダーは、壁面および天井の二箇所に設置し、人物に対して異なる視点からの反射情報を取得できるよう構成した (図 5.1 参照)。

- 壁面設置: 床面から高さ約 1.2m の位置

- 天井設置: 壁面から約 2.0m 離れた天井中央位置

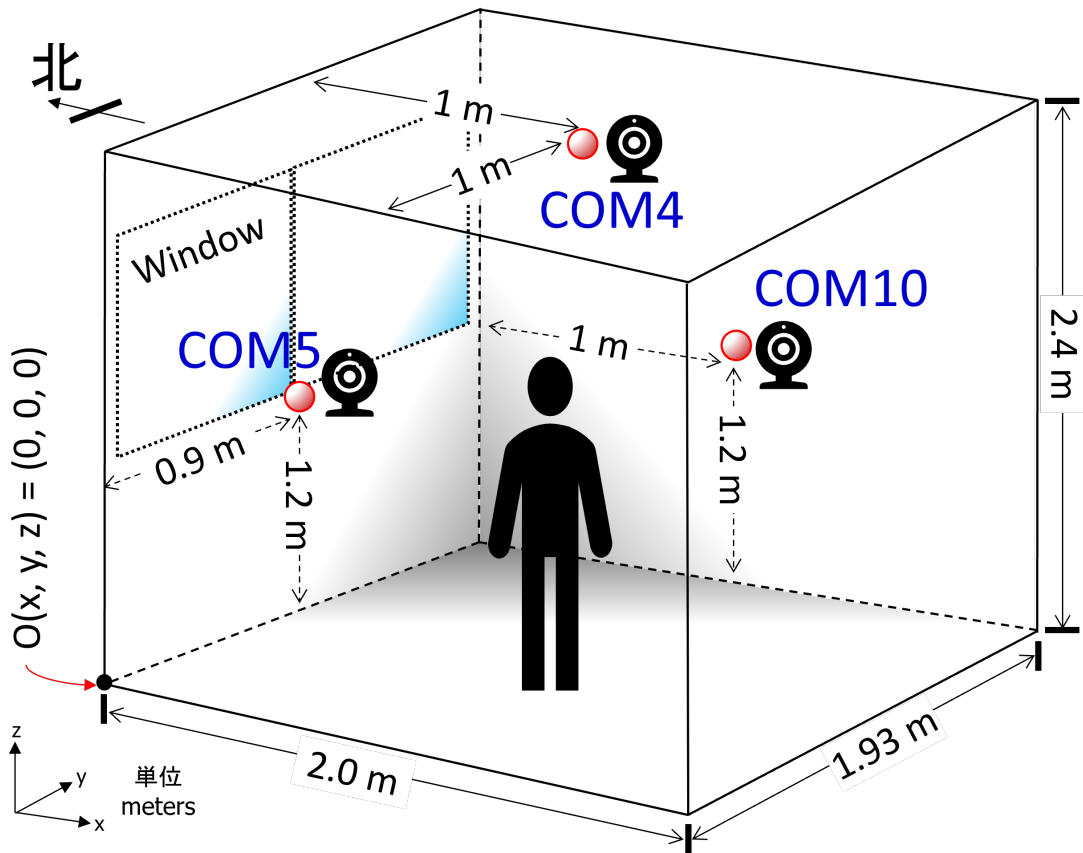


図 5.1: 実験環境およびセンサ配置図

このような直交に近い配置により、単一センサでは遮蔽（オクルージョン）により捉えにくい姿勢変化や、センサに対して正対・並行する動作の違いを相補的に取得することが可能となる。なお、実験ではセンサデータのラベリングおよび挙動確認のため、同時にカメラによる撮影を行った。

## 5.2 姿勢動作の計測

実環境における人間活動は多岐にわたるが、本研究では先行研究 [9] で定義された代表的な日常動作を基準として実験シナリオを構築した。先行研究で定義される 6 動作のうち、立位状態での腕の動作要素が強い「手を振る (Waving)」は本研究の主目的（姿勢変化の推定）と性質が異なるため除外し、姿勢遷移に焦点を当てた以下の 5 動作を計測対象とした。

計測対象とした。各動作の定義および状態遷移の対応を以下に示す。

1. **Standing up** (Sit → Stand) : 座位姿勢から立位姿勢へ移行する動作。

2. **Sitting down** (Stand → Sit) : 立位姿勢から座位姿勢へ移行する動作.
3. **Sleeping** (Stand / Sit → Sleep) : 立位姿勢からベットで仰向けで寝ている姿勢へ移行し、その状態を維持する動作.
4. **Falling** (Stand / Walking → Fall) : 立位姿勢から急激に床に伏せる (転倒状態) へ移行する動作.
5. **Getting up** (Sleep → Sit / Stand) : ベットで仰向けで寝ている姿勢から上半身を起こし、立位姿勢へ移行する動作.

被験者5名を対象に, 上記5動作についてセンサに対する体の向き (方位) を東・西・南・北の4方向に変えながら, 各方向10回ずつ反復計測を行った. 収集したデータ内訳を表5.1に示す.

表 5.1: 収集した姿勢動作データセットの詳細

姿勢シナリオ	動作方向	回数/方向	被験者数	総サンプル数
Standing up	4	10	5	200
Sitting down	4	10	5	200
Sleeping	4	10	5	200
Getting up	4	10	5	200
Falling	4	10	5	200
<b>合計</b>	-	-	<b>5</b>	<b>1000</b>

センサのサンプリングレートは10 Hz (0.1 秒間隔) とし, 1 試行あたり 10 秒間 (100 フレーム) のデータを取得した. したがって, 各サンプルは時間長  $T = 100$  の時系列データとして構成される.

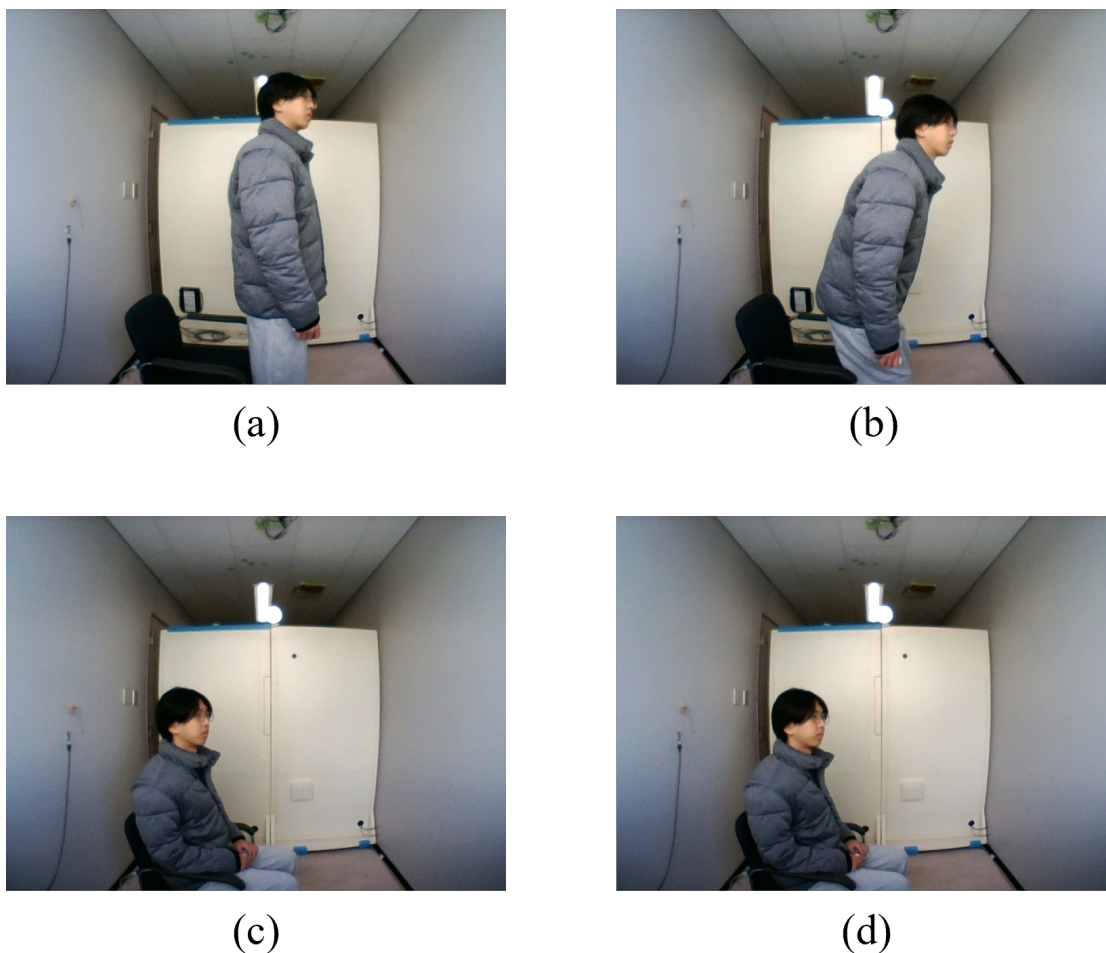


図 5.2: 姿勢動作計測の様子 (a) 初期化,(b) 動作,(c) 予備,(d) 終了

### 5.2.1 データセットの構成

本研究では, 取得した時系列データを以下の3種類に構成し, モデルの学習および評価に用いた.

#### 単独動作データ (1P)

単独動作データ (1P) は, 単一の姿勢動作を1試行 (10秒, 100フレーム) として記録した基本単位である. 各サンプルは初期状態から動作を経て最終状態へ移行する過程を含む. 本研究では合計1,000サンプルを収集した (表5.1). また, 計測の安定性を確保するため, 各10秒の試行は「初期化 (2秒)」「動作 (3秒)」「予備 (3秒)」「終了処理 (2秒)」の計10秒で構成した. なお, 1Pでは取得した100フレームをそのままモデル入力として用い, 単独動作の識別性能を評価する.

## 組み合わせ動作データ (2P)

組み合わせ動作データ (2P) は, 単独動作データ (1P) を 2 つ時間方向に連結することで, 連続的な姿勢遷移を疑似的に構成した評価用データである. 具体的には, 異なる 2 動作のサンプル (各 100 フレーム) を連結し, 全長 200 フレームの系列として扱う. 例として日常的な姿勢遷移として代表的な以下の 3 パターンを示す.

- **Sitting to Sleeping:** Sitting down → Sleeping
- **Sitting to Standing:** Sitting down → Standing up
- **Sleeping to Standing:** Sleeping → Standing up

なお, 2P は計測により新規に取得したデータではなく, 収集済みの 1P を連結して生成した系列である. 探索によって前後の重複を許容する全組み合わせを生成する. この 2P 系列に対してスライディングウィンドウを適用し, 動作遷移前後における識別の安定性を評価する.

## 長尺連続動作データ (10P)

より複雑な日常シーンを想定し, 複数の動作要素を連続して行う長尺動作系列 (10P) を構築した. 10P は複数の動作要素を時間方向に連結して生成した系列であり, 長時間の文脈を含む連続動作に対して, モデルが過去の情報を保持しつつ姿勢変化へ追従できるかを評価する.

## 5.3 姿勢推定モデルの構築

本節では, 収集したレーダーデータを用いて姿勢動作の分類を行うためのモデル構築について述べる. 本研究では, 時系列データに内在する空間的・時間的依存関係を同時に学習可能な時空間グラフ畳み込みネットワーク (Spatial-Temporal Graph Convolutional Network: ST-GCN) をベースラインとして採用し, これにセンサ信頼度に基づく動的グラフ構築機構を導入した. アンカー選択に基づく時空間グラフ畳み込みネットワーク (ST-GCN/AG) との比較を行う.

### 5.3.1 データ前処理

収集されたレーダーデータは, 複数センサから得られる距離ゲートごとのエネルギー情報を時系列として保持している. 本研究では, 使用するセンサ数を  $V = 3$ , 各サンプルに含まれるフレーム数を  $T = 100$  と設定した.

各フレームにおいて,1 センサあたり 18 次元の特徴量 (距離ゾーンごとの動体・静止体エネルギーなど) を抽出し, これを入力チャンネル数  $C = 18$  と定義する. したがって, ネットワークへの入力は

$$(N, C, T, V)$$

の 4 次元テンソルとして構成される. ここで  $N$  はバッチサイズを表す.

学習時には, データ正規化としてノード次元とチャンネル次元を統合したバッチ正規化を適用し, センサ間スケール差の影響を低減している.

### 5.3.2 ST-GCN/SA による動的グラフ学習

ST-GCN/SA では, 複数の準ミリ波レーダーセンサから得られる時系列信号を, ノード間の関係性を含む時空間グラフとして表現し, その関係構造を空間アテンションによって動的に再構成する.

本研究においては, 各センサをグラフのノードとし, 各時刻におけるセンサ出力をノード特徴として扱う. 入力データは, チャンネル数  $C$ , 時間長  $T$ , ノード数  $V$  からなるテンソル  $X \in \mathbb{R}^{C \times T \times V}$  として表現される. ここで, 各ノードは単一センサに対応し, その特徴ベクトルは距離ゲートごとのエネルギー分布など, センサが観測した空間情報を含んでいる.

従来の ST-GCN では, センサ間の関係はあらかじめ定義された静的隣接行列によって固定されるが, ST-GCN/SA では入力データに基づいてノード間の関係性を動的に更新する. 具体的には, 時間方向の平均により, 各ノードの時間的変動を集約した空間特徴  $X_{\text{spatial}}$  を生成する. この操作により, 一時的なノイズや瞬間的な変動を抑制し, 各センサの相対的な寄与関係を反映した表現が得られる.

次に, この空間特徴に基づいてノード間の類似度を算出し, Softmax 正規化により動的隣接行列  $A_{\text{attn}}$  を構築する. この行列は, 各ノードが他のノードに対してどの程度の重要度を持つかを表し, センサ間の関係性をデータ駆動的に表現するものである.

このようにして得られた動的隣接行列は, 静的隣接行列と組み合わせて用いられ, 各時系列データに応じて変化するセンサ間の依存関係を反映する. その結果, センサ配置や被験者の位置関係の変化に応じて, 重要なセンサ間の結合が強調され, 固定的なグラフ構造のみでは捉えにくい空間的特徴を表現することが可能となる.

以上より, ST-GCN/SA におけるセンサデータの表現は, 各センサの観測値をノード特徴としてグラフ上に配置し, その関係構造を入力データに基づいて動的に再構成することで, 柔軟な空間モデリングを実現するものである.

### 5.3.3 ST-GCN/AG による動的グラフ学習

提案手法である ST-GCN/AG では、複数の準ミリ波レーダーセンサから得られる時系列信号を時空間グラフとして表現する際に、入力データの信頼度に基づいてノード間の関係性を動的に再構成するアンカー誘導型グラフを導入する。

本研究では、各センサをノードとし、各時刻におけるセンサ出力をノード特徴とすることで、入力データを  $X \in \mathbb{R}^{C \times T \times V}$  の形式で表現する。ここで、各ノードの特徴は、距離ゲートごとのエネルギー分布など、センサが観測した空間情報を含んでいる。

まず、各ノードの信頼度を推定するため、時間方向における移動分散を用いて、センサ出力の時間的変動を評価する。この指標に基づき、各ノードに対して信頼度  $S$  を割り当てることで、各時刻におけるセンサの寄与度を定量化する。この信頼度は正規化され、ノードごとの重要度として解釈される。

次に、信頼度に基づいて代表的な特徴を持つアンカーを生成する。具体的には、信頼度を重みとして各ノード特徴を統合することで、信頼度の高いセンサを中心としたアンカー特徴を構成する。その後、このアンカー特徴と各ノード特徴との類似度を計算することで、ノード間の関係性を表す動的隣接行列  $A_{\text{anchor}}$  を構築する。

この動的隣接行列は、静的隣接行列と組み合わせて用いられ、信頼度の高いセンサを起点とした情報伝搬を強調する役割を持つ。これにより、センサ間の関係性は単なる類似度に基づくのではなく、信頼度という物理的関連性に基づいて制約された形で再構成される。

さらに、信頼度に基づくゲーティング機構を導入することで、信頼度の低いセンサやフレームに由来する特徴量の影響を抑制する。この処理により、ノイズを多く含むセンサ環境においても、重要な情報を保持したまま安定した特徴表現を獲得することが可能となる。

以上より、ST-GCN/AG では、センサデータをノード特徴としてグラフ上に配置し、信頼度に基づいて関係構造を動的に誘導することで、入力データに応じた柔軟かつ制約付きの空間モデリングを実現する。

### 5.3.4 学習条件

ST-GCN および ST-GCN/SA, ST-GCN/AG の各モデルに対して、以下の共通条件のもとで学習を行った。

- エポック数：100
- 学習率： $5.0 \times 10^{-4}$  (ST-GCN) ,  $1.0 \times 10^{-4}$  (ST-GCN/AG)
- 最適化手法：Adam (Adaptive Moment Estimation)
- 正則化 (Weight Decay)： $1.0 \times 10^{-7}$

- 損失関数：交差エントロピー誤差（Cross Entropy Loss）

学習は誤差逆伝播法により行い、各エポック終了時に検証データに対する認識精度を算出した。検証精度が最大となった時点のモデルを最良モデルとして保存し、以降の評価ではこのモデルを用いた。

# 第6章 評価

## 6.1 はじめに

本節では、構築した姿勢推定モデルの性能を定量的に評価する。評価指標として、各動作クラスに対する正解率（Accuracy）を用いる。Accuracy は、全テストサンプル数に対する正しく分類されたサンプル数の割合として定義する。また、クラスごとの誤分類傾向を詳細に分析するため、正解クラスごとに正規化した混同行列（Confusion Matrix）を作成する。

比較対象として、ベースライン手法である従来の ST-GCN と、本研究で提案する ST-GCN/SA と ST-GCN/AG の認識精度を比較検討する。評価は Python および PyTorch を用いて実装し、学習済みモデルに対してテストデータを適用することで実施した。

## 6.2 評価手法

### 6.2.1 学習過程の評価

提案モデルの学習における収束性および汎化性能を評価するため、学習曲線（Loss Curve）の推移を分析する。

本実験では、エポック数を 100 に設定し、損失関数にはクロスエントロピー誤差を用いた。学習データ（Training）および検証データ（Validation）に対する損失の推移を比較することで、モデルの収束挙動および過学習の有無を評価する。

また、提案手法であるアンカー誘導機構が、学習の安定性および収束速度に与える影響について、ベースライン手法との比較を通じて検討する。

### 6.2.2 単独動作および動作遷移系列に対する評価

#### 単独姿勢動作データの評価

単一の姿勢動作のみを含む 100 フレームの時系列データを用い、5 種類の姿勢動作（Standing Up, Sitting Down, Sleeping, Falling, Getting Up）に対する分類性能を評価した。

評価にはテストデータを用い、各サンプルに対する予測ラベルと正解ラベルの一致率として Accuracy を算出した。

さらに、Falling と Sleeping、および Sitting Down と Sleeping のように、重心位置が類似する動作間の識別性能に着目し、混同行列を用いて誤分類傾向を分析する。

### 姿勢動作遷移データの評価

次に、動作遷移を含む系列に対する認識安定性を評価する。実環境では動作は連続的に発生するため、動作遷移中においても安定した認識が可能であることが求められる。

本実験では、2つの異なる動作 ( $A, B$ ) を時間方向に結合した 200 フレームの系列データを生成し、スライディングウィンドウ法により評価を行った。ストライド幅 2 で 100 フレームのウィンドウを適用し、動作遷移を含む疑似的な時系列データを構築した。

評価に際しては、遷移過渡期の曖昧な状態を除外するため、以下の安定区間における正答率 (Accuracy) を指標とした。

- **安定区間 A**：開始位置が 0～25 フレームの範囲において、動作  $A$  を正しく分類した割合。
- **安定区間 B**：開始位置が 75 フレーム以降の範囲において、動作  $B$  を正しく分類した割合。

これにより、動作遷移を含む状況において、モデルが前後の動作を安定して識別できるかを評価する。

### 6.2.3 ネットワーク構成およびデータ分割条件の比較設定

提案モデルの頑健性および最適構成を明らかにするため、ネットワークの深さ (層数) および学習データの分割比率を変更した条件下で評価を行った。

#### 畳み込み層の深さ

基本モデルでは 5 層の ST-GCN ブロック ( $l_1 \sim l_5$ ) を採用しているが、以下の 3 通りの構成に変更し、認識精度を比較した。

- **3 層構成**： $l_1 \sim l_3$  のみを使用した最小構成。
- **5 層構成 (基準)**：性能と計算効率のバランスを考慮した構成。
- **9 層構成**：長期的な時間依存性の学習を目的とした構成。

## データ分割比率

学習データ量の違いが汎化性能に与える影響を調査するため、学習データ割合を50%,33%,25%とした3種類の分割比率を設定した。各条件におけるデータの内訳は以下の通りである。

**学習データ 50%** : 学習データ 50%, 検証データ 25%, テストデータ 25%

**学習データ 33%** : 学習データ 33%, 検証データ 33%, テストデータ 33%

**学習データ 25%** : 学習データ 25%, 検証データ 37.5%, テストデータ 37.5%

これらの条件において得られた認識精度を比較することで、学習データ量およびネットワーク層数が姿勢動作認識性能に与える影響を分析する。

## 6.3 結果

本章では、前章で述べた評価手法に基づき、提案手法である ST-GCN/SA と ST-GCN/AG、ベースライン手法である ST-GCN の姿勢推定性能に関する実験結果を示す。

評価は、学習過程における損失の推移、単独姿勢動作および単独姿勢動作データを組み合わせた場合の連続した姿勢推定動作遷移に対する分類精度、ならびにネットワーク構成および学習データ分割条件の違いによる性能変化の三つの観点から行った。

本章では、まず畳み込み層の深さによる影響を示す。次に畳み込み層の深さを基に学習データ量の違いがモデルの収束性および汎化性能に与える影響を明確にするため、データ分割比率の異なる3つの条件（パターン A,B,C）を設定し、それぞれについて詳細な比較評価を行う。

### 畳み込み層の深さによる影響

ネットワークの表現能力を決定する重要な要因である ST-GCN ブロック（畳み込み層）の積層数による精度の変化を調査した。本実験では、層数を3層,5層,9層に設定し、モデル（ST-GCN,ST-GCN/AG,ST-GCN/SA）およびデータ条件（単独,2P,10P）ごとの全体精度（Overall Accuracy）を比較した。

表 6.1 に、層数および手法ごとの全体精度をまとめる。

表 6.1: ネットワーク層数および手法別の姿勢動作認識精度

層数	手法	1P	2P	10P
3層	ST-GCN	<b>97.00 %</b>	<b>97.25 %</b>	<b>97.35%</b>
	ST-GCN/ST	91.50%	92.01%	91.82%
	ST-GCN/AG	88.50%	88.83%	86.44%
5層	ST-GCN	89.50%	85.93%	87.58%
	ST-GCN/ST	88.50%	85.12%	84.43%
	ST-GCN/AG	<b>93.00%</b>	<b>91.05%</b>	<b>89.43%</b>
9層	ST-GCN	90.00%	86.92%	86.95%
	ST-GCN/ST	89.50%	84.49%	83.23%
	ST-GCN/AG	<b>93.50%</b>	<b>91.30%</b>	<b>91.75%</b>

結果を図 6.1～図 6.3 および表 6.1 に示す.

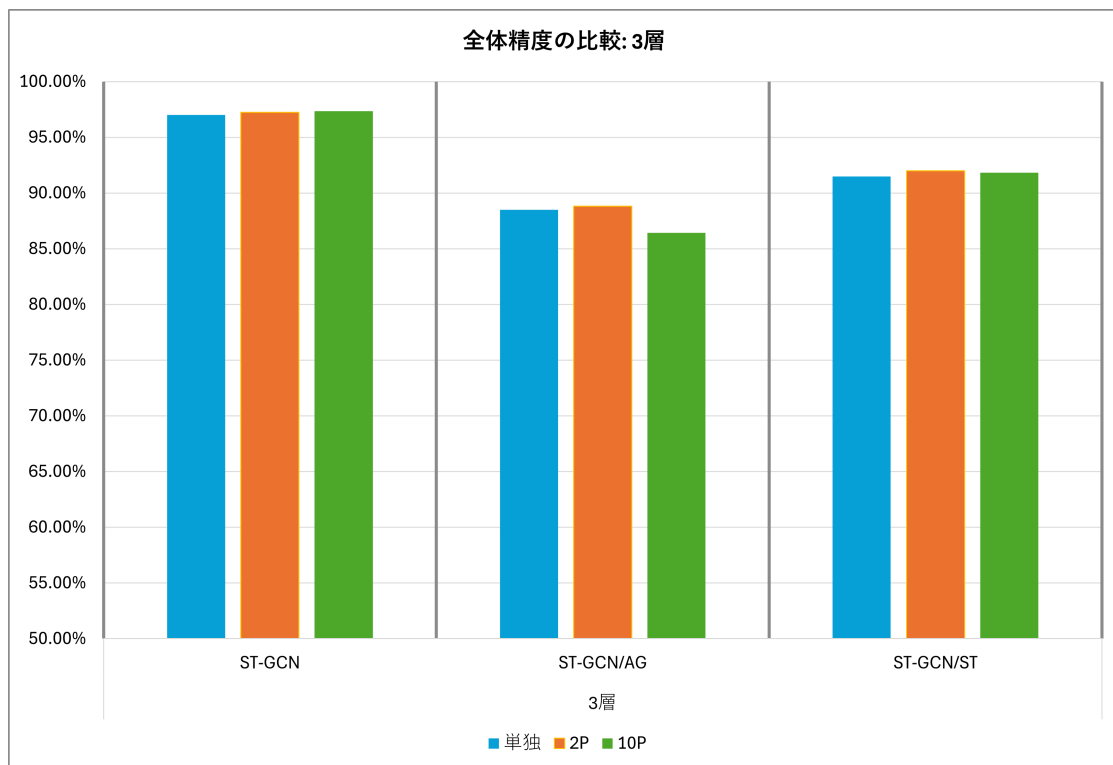


図 6.1: 全体精度: 3層

次に, 図 6.2 に示すように, 5層構成では ST-GCN/AG が全条件において最も高い精度を示した. 単独動作で約 93%, 2P で約 91%, 10P で約 89–90% であった. 一方で, ST-GCN は単独動作で約 89.50%, 2P で約 85.93%, 10P で約 87.58%, ST-GCN/SA は単

独動作で約 88.50%,2P で約 85.12%,10P で約 84.43%となり,いずれも ST-GCN/AG を下回る結果となった.

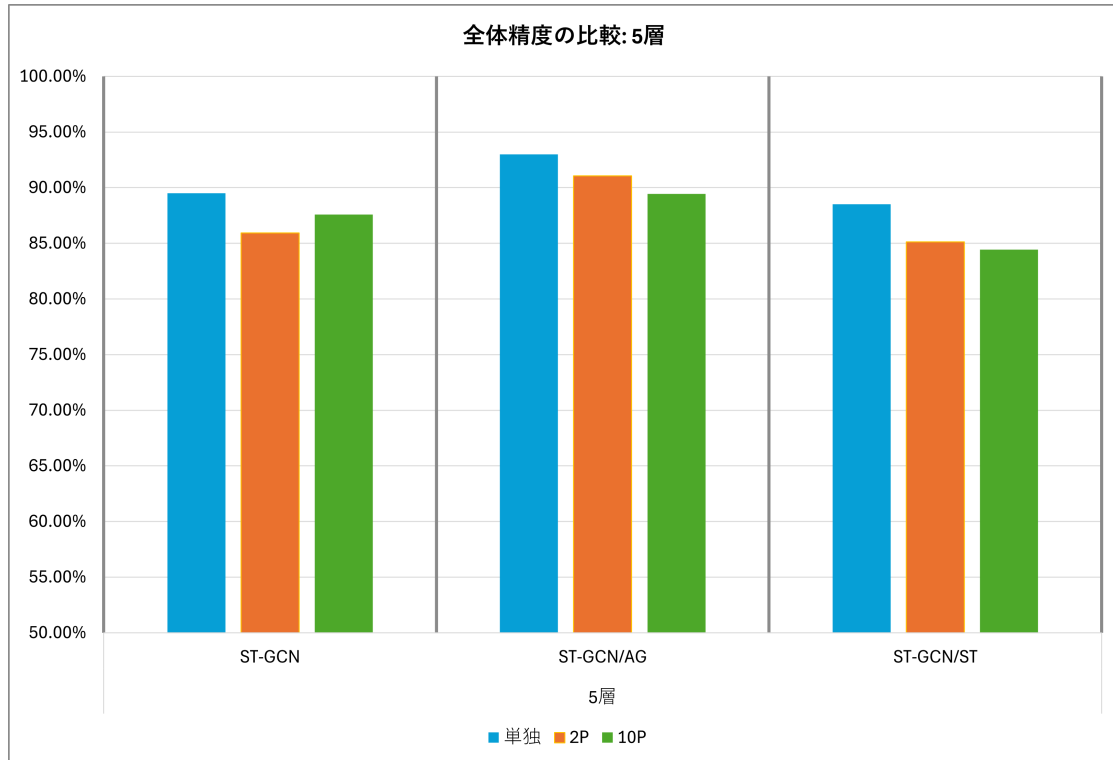


図 6.2: 全体精度: 5層

さらに, 図 6.3 に示すように,9層構成では ST-GCN/AG が引き続き最も高い精度を示した. 具体的には, 単独動作で約 93-94%,2P で約 91%,10P で約 92%であった.

特に,10P 条件では 5層構成と比較して精度の向上が確認された. 一方で,ST-GCN は単独動作で約 90.00%,2P で約 86.92%,10P で約 86.95%,ST-GCN/SA は単独動作で約 89.50%,2P で約 84.49%,10P で約 83.23%となり,いずれも 5層構成から大きな改善は見られなかった. また,2P および 10P 条件では精度の低下が確認された.

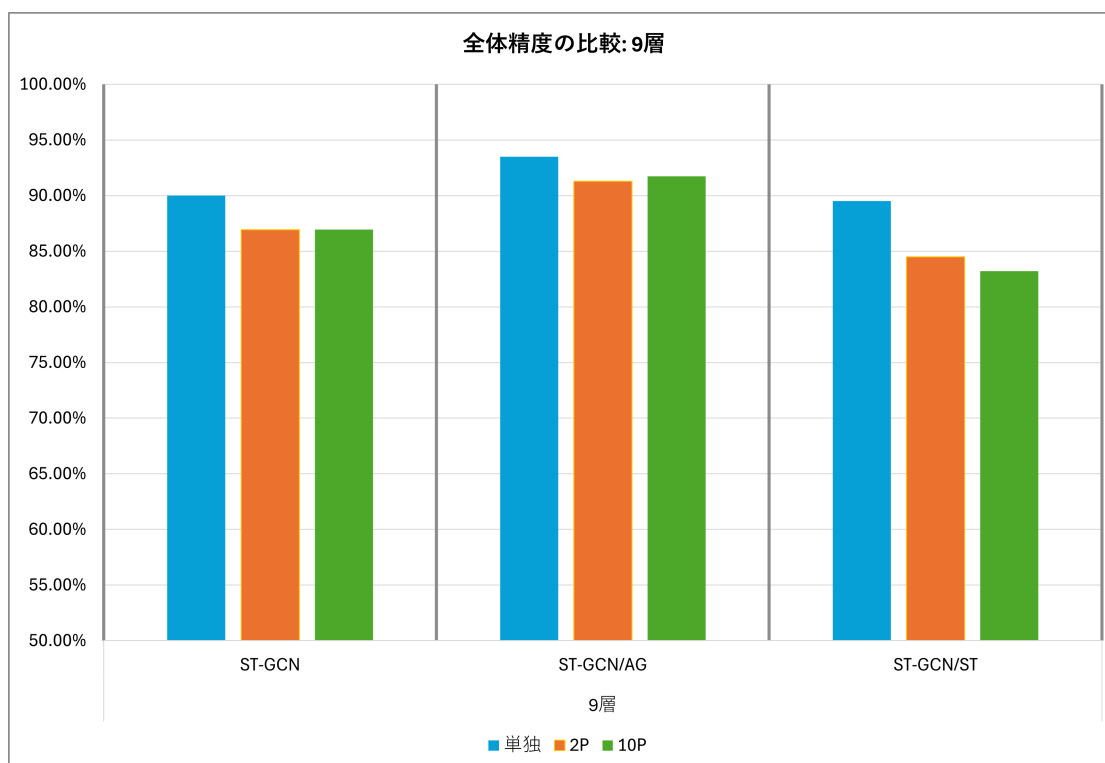


図 6.3: 全体精度: 9層

畳み込み層の深さを3層、5層、9層と変化させた際の認識精度を検討した。その結果、層の深さがモデルの認識性能に影響を与えることが確認された。手法別の傾向として、ST-GCN/AGは5層および9層の構成において一貫して高い精度を示した。一方で、9層構成は一部の条件下で最高精度を記録したものの、5層構成に対する精度向上幅は限定的であった。これらの結果に基づき、本研究ではモデルの精度と計算量・パラメータ数のバランス、および各提案手法（ST-GCN/SAおよびST-GCN/AG）の性能差が明確に現れる点を考慮し、以降の比較評価には5層モデルを採用する。

### 6.3.1 データ分割比率による精度の比較

次に、学習データ量の違いが認識性能に与える影響を調査するため、5層構成のネットワークを用い、学習データ割合を50%、33%、25%とした3種類の分割比率について評価を行った。各条件において、単独動作データ（1P）、組み合わせ動作データ（2P）、長尺連続動作データ（10P）に対する全体の認識精度を比較した。

学習データ 50%

表 6.2: 単独姿勢推定における精度 (学習データ 50%)

姿勢動作	ST-GCN	ST-GCN/AG	ST-GCN/SA
Sitting Down	90.00%	<b>98.00%</b>	96.00%
Standing Up	98.00%	<b>100.00%</b>	99.00%
Sleeping	<b>99.50%</b>	94.50%	92.50%
Getting Up	94.00%	<b>98.50%</b>	96.00%
Falling	<b>97.50%</b>	95.00%	93.50%
<b>Overall</b>	89.50%	<b>93.00%</b>	88.50%

表 6.3: 組み合わせ姿勢動作 (2 組) における精度 (学習データ 50%)

姿勢動作	ST-GCN	ST-GCN/AG	ST-GCN/SA
Sitting Down	89.09%	<b>96.75%</b>	93.96%
Standing Up	97.28%	<b>99.63%</b>	98.31%
Sleeping	<b>97.11%</b>	93.38%	90.77%
Getting Up	92.82%	<b>99.09%</b>	95.49%
Falling	<b>95.56%</b>	93.25%	91.71%
<b>Overall</b>	85.93%	<b>91.05%</b>	85.12%

表 6.4: 組み合わせ姿勢動作 (10 組) における精度 (学習データ 50%)

姿勢動作	ST-GCN	ST-GCN/AG	ST-GCN/SA
Sitting Down	90.32%	<b>94.77%</b>	93.64%
Standing Up	97.47%	<b>98.02%</b>	98.27%
Sleeping	<b>97.76%</b>	93.77%	90.32%
Getting Up	93.43%	<b>97.90%</b>	95.36%
Falling	<b>96.17%</b>	94.40%	91.27%
<b>Overall</b>	87.58%	<b>89.43%</b>	84.43%

学習データ 33%

表 6.5: 単独姿勢推定における精度 (学習データ 33%)

姿勢動作	ST-GCN	ST-GCN/AG	ST-GCN/SA
Sitting Down	93.32%	<b>97.17%</b>	94.60%
Standing Up	90.75%	<b>98.97%</b>	90.49%
Sleeping	81.49%	<b>94.09%</b>	92.54%
Getting Up	87.92%	94.86%	<b>97.43%</b>
Falling	93.57%	90.23%	<b>94.60%</b>
<b>Overall</b>	73.52%	<b>87.66%</b>	84.83%

表 6.6: 組み合わせ姿勢動作 (2 組) における精度 (学習データ 33%)

姿勢動作	ST-GCN	ST-GCN/AG	ST-GCN/SA
Sitting Down	93.16%	<b>95.67%</b>	93.92%
Standing Up	89.98%	<b>98.44%</b>	90.85%
Sleeping	79.06%	<b>92.71%</b>	90.19%
Getting Up	87.36%	94.53%	<b>96.54%</b>
Falling	<b>91.84%</b>	87.18%	92.45%
<b>Overall</b>	70.70%	<b>84.27%</b>	81.98%

表 6.7: 組み合わせ姿勢動作 (10 組) における精度 (学習データ 33%)

姿勢動作	ST-GCN	ST-GCN/AG	ST-GCN/SA
Sitting Down	93.56%	<b>94.17%</b>	93.74%
Standing Up	90.05%	<b>92.58%</b>	90.78%
Sleeping	79.56%	<b>93.84%</b>	89.65%
Getting Up	87.59%	91.98%	<b>96.43%</b>
Falling	<b>92.02%</b>	91.65%	91.97%
<b>Overall</b>	71.39%	<b>82.11%</b>	81.29%

## 学習データ 25%

表 6.8: 単独姿勢推定における精度 (学習データ 25%)

姿勢動作	ST-GCN	ST-GCN/AG	ST-GCN/SA
Sitting Down	93.32%	<b>98.20%</b>	93.32%
Standing Up	88.17%	<b>100.00%</b>	88.17%
Sleeping	87.15%	<b>92.80%</b>	87.15%
Getting Up	95.37%	<b>97.94%</b>	95.37%
Falling	90.75%	<b>93.06%</b>	90.75%
<b>Overall</b>	77.38%	<b>91.00%</b>	77.38%

表 6.9: 組み合わせ姿勢動作 (2 組) における精度 (学習データ 25%)

姿勢動作	ST-GCN	ST-GCN/AG	ST-GCN/SA
Sitting Down	93.21%	<b>97.73%</b>	93.21%
Standing Up	87.38%	<b>99.75%</b>	87.38%
Sleeping	84.95%	<b>91.87%</b>	84.95%
Getting Up	95.66%	<b>97.89%</b>	95.66%
Falling	88.34%	<b>91.91%</b>	88.34%
<b>Overall</b>	74.77%	<b>89.57%</b>	74.77%

表 6.10: 組み合わせ姿勢動作 (10 組) における精度 (学習データ 25%)

姿勢動作	ST-GCN	ST-GCN/AG	ST-GCN/SA
Sitting Down	92.96%	<b>96.42%</b>	92.96%
Standing Up	87.33%	<b>99.38%</b>	87.33%
Sleeping	84.65%	<b>93.19%</b>	84.65%
Getting Up	95.63%	<b>97.07%</b>	95.63%
Falling	87.68%	<b>91.30%</b>	87.68%
<b>Overall</b>	74.12%	<b>88.68%</b>	74.12%

## 各パターンにおける全体精度の比較

5層構成に固定し, 学習データ割合を 50%, 33%, 25% に変化させて比較した結果, ST-GCN/AG はいずれの条件においても ST-GCN を上回る性能を示した. 特に 33% お

表 6.11: 学習データ分割比率別における全体認識精度

学習データ割合	動作種別	ST-GCN	ST-GCN/SA	ST-GCN/AG
50%	1P	89.50%	88.50%	<b>93.00%</b>
	2P	85.93%	85.12%	<b>91.05%</b>
	10P	87.58%	84.43%	<b>89.43%</b>
33%	1P	73.52%	84.83%	<b>87.66%</b>
	2P	70.70%	81.98%	<b>84.27%</b>
	10P	71.39%	81.29%	<b>82.11%</b>
25%	1P	77.38%	77.38%	<b>91.00%</b>
	2P	74.77%	74.77%	<b>89.57%</b>
	10P	74.12%	74.12%	<b>88.68%</b>

よび 25%の少量データ条件においても性能低下が小さく、高い安定性を維持することが確認された。これは、各フレームで推定したノード信頼度に基づいてアンカーを決定し、そのアンカーを起点としてセンサ間の空間関係を誘導することで、学習時における空間関係推定の不確かさが抑制されるためであると考えられる。一方、ST-GCN/SA は 33%条件では ST-GCN を上回るものの、25%条件では単独動作の分類精度が低下し、少量データ環境下での安定性に課題が残る。以上より、5 層構成においては、ST-GCN/AG は学習データ量への依存性が小さく、データ取得に制約のあるセンサベース姿勢推定において有効な手法であることが示された。

## 6.4 考察

本研究では、ST-GCN、ST-GCN/SA、ST-GCN/AG の各モデルについて、ネットワークの層数を変化させた際の認識性能を比較した。その結果、層の深さとモデル構造の違いが性能に大きく影響することが明らかとなった。

まず 3 層構成では、図 6.1 に示すように、ST-GCN が最も高い精度を示した。これは、浅いネットワークにおいては、固定グラフに基づく単純な時空間特徴抽出のみでも十分に識別が可能であるためと考えられる。一方で、ST-GCN/AG および ST-GCN/SA では、動的な関係性を導入するための追加パラメータが存在するが、浅い構造ではそれらを十分に活用するための表現能力が不足し、結果として性能向上に寄与しなかったと考えられる。

次に 5 層構成では、図 6.2 に示すように、ST-GCN/AG が最も高い精度を示した。これは、層を深くすることで、フレームごとに推定される信頼度に基づく空間集約（アンカー誘導）が段階的に反映されるためであると考えられる。また、時間方向の畳み込みにより動作遷移のパターンが抽象化され、動作ごとに変化するセンサ間関係をより適切に捉えられるようになったと推察される。

一方で,ST-GCN/SA は全条件でST-GCN/AG を下回った (図6.2) . この結果は, 単純なアテンション機構による関係性の学習では, ノイズを含むセンサデータにおいて安定した関係推定が難しい可能性を示唆している. これに対し,ST-GCN/AG では, 信頼度に基づいてアンカーを明示的に選択し, 空間関係を制御する構造を有しているため, より安定した特徴抽出が可能であったと考えられる.

さらに9層構成においても, 図6.3 に示すように, ST-GCN/AG が最も高い精度を維持した. これは, 層を重ねることで, 信頼度に基づくアンカー決定を介した空間関係の制御がより多段的に反映されるためと考えられる. 一方で,ST-GCN およびST-GCN/SA では, 層を深くしても性能の向上は限定的であり, 特に2P および10P 条件では精度の低下が確認された (図6.3) . このことから, 固定的なグラフ構造や単純なアテンションのみでは, 動作に応じて変化するセンサ間依存を安定に捉え続けることが困難である可能性が示唆される.

また, 表6.1 に示すように, ST-GCN/AG は5層および9層において一貫して高い精度を示しており, 層数の増加に対して安定した性能を維持していることが確認できる.

以上の結果より, 動的な空間関係を扱う際には, 単に関係性を学習するだけでなく, その推定過程に対して適切な制約や指針を与えることが重要であると考えられる. 特に, 本研究で提案したアンカー誘導機構は, ノイズを含むセンサ環境においても安定した関係推定を可能とし, 深層化した場合でも性能を維持できる有効な手法であるといえる.

## 6.5 まとめ

本節では, ネットワークを5層構成に固定した上で, 学習データ割合を50%,33%,25%に変化させ, ST-GCN,ST-GCN/SA,ST-GCN/AG の性能を比較した.

その結果,ST-GCN/AG は全てのデータ割合条件においてST-GCN を上回る精度を示した. また,33%および25%の少量データ条件においても, 高い精度を維持しており, データ量の減少に対して安定した性能を示すことが確認された.

一方で,ST-GCN/SA は33%条件ではST-GCN を上回るものの, 25%条件では精度の低下が見られ, 少量データ環境下での安定性に課題があることが示された.

以上より,5層構成においては, ST-GCN/AG は学習データ量への依存性が小さく, データ取得に制約のあるセンサベース姿勢推定において有効な手法であることが明らかとなった.

# 第7章 おわりに

## 7.1 本研究の概要と成果

本研究では、24 GHz 帯の準ミリ波レーダーを用いた非接触かつプライバシー配慮型の姿勢推定を目的として、複数センサから得られる時系列信号を時空間グラフとして表現し、姿勢動作の分類を行う手法を提案した。近年、カメラを用いた姿勢推定手法は高精度化が進んでいる一方で、照明条件や遮蔽の影響、およびプライバシー保護の観点から、実環境への適用には依然として課題が残されている。これに対し、本研究ではレーダーセンサを用いることで、環境光の影響を受けにくく、かつ個人情報を直接取得しない形での姿勢推定の実現を目指した。

提案手法は、実環境におけるセンサ配置および計測に基づくデータ収集（ハードウェアフェーズ）と、センサ特性を考慮した補正・前処理および学習モデルの構築（ソフトウェアフェーズ）から構成される。ハードウェアフェーズでは、複数の準ミリ波レーダーを用いて、日常的な姿勢動作（Standing Up, Sitting Down, Sleeping, Falling, Getting Up）を対象としたデータセットを構築した。ソフトウェアフェーズでは、各センサから得られる時系列信号をノードとし、センサ間の関係をエッジとして表現することで、入力データを時空間グラフとして定式化した。

学習モデルとしては、時系列データの空間・時間依存を同時に扱うことが可能な Spatial-Temporal Graph Convolutional Network (ST-GCN) をベースラインとした。さらに、センサ信頼度に基づいてアンカーを動的に選択し、そのアンカーを起点としてノード間の関係性を誘導する手法、すなわちアンカー選択に基づく時空間グラフ畳み込みネットワーク (ST-GCN/AG) を提案した。本手法では、各フレームにおけるセンサ出力の安定性から信頼度を推定し、信頼度の高いノードをアンカーとして選択することで、ノイズの影響を受けにくい関係構造を構築する。この関係推定は、ノード間の相関を重みとして表現する点で Attention 機構と類似しているが、信頼度に基づく明示的な制約を導入している点に特徴がある。これにより、データ駆動型の Attention に比べて、センサノイズや環境変動に対して安定した関係推定が可能となる。

評価実験では、単独動作（1P）および動作遷移を含む連続動作（2P, 10P）に対する分類精度を指標とし、さらにネットワークの深さ（3層, 5層, 9層）および学習データ割合（50%, 33%, 25%）の違いによる性能変化を比較した。その結果、浅い3層構成では固定グラフに基づく ST-GCN が高い精度を示した一方で、層を深くすることで ST-GCN/AG の性能が顕著に向上し、5層および9層構成において最も

高い認識精度を達成した。特に、動作遷移を含む複雑な系列（2P,10P）において、ST-GCN/AGは安定した識別性能を示し、従来手法に対して優位性を有することが確認された。

また、学習データ量を制限した条件においても、ST-GCN/AGは一貫してST-GCNを上回る精度を示し、特に33%および25%の少量データ条件においても性能低下が抑制されることを確認した。この結果は、アンカー誘導により信頼度の高いセンサ情報を優先的に活用することで、空間関係推定の不確かさが低減され、限られたデータ環境においても安定した学習が可能となることを示している。一方で、空間アテンションを導入したST-GCN/SAは、中程度のデータ量では一定の性能向上を示すものの、少量データ条件では不安定となる傾向が確認された。

以上の結果より、本研究で提案したST-GCN/AGは、分解能の低い準ミリ波レーダーを用いた姿勢推定において、センサ配置差やノイズの影響を低減しつつ、安定した姿勢認識を実現可能な手法であることが明らかとなった。特に、データ取得に制約のある実環境においても高い汎化性能を維持できる点は、実用的なセンサベース姿勢推定システムの構築において重要な知見である。

本研究は、レーダーセンサを用いた姿勢推定において、空間関係の動的モデリングに信頼度という物理的関連性を導入することで、従来のグラフ畳み込み手法を拡張した点に特徴がある。このアプローチにより、ノイズを含む実環境データに対しても安定した特徴の学習が可能となり、今後のセンサ融合や行動理解への応用においても有用な基盤となることが期待される。

## 7.2 今後の課題

本研究の結果を踏まえ、今後の課題として以下の点が挙げられる。

第一に、動作遷移の過渡区間における認識の安定化である。本研究では、動作遷移に伴う曖昧な状態を避けるため、安定区間に限定した評価を行った。しかしながら、実環境においては、動作の切り替わり直後に正確に検知できるかが重要となる。特に、転倒検知などの応用においては、遷移過程そのものの識別性能が求められる。このため、過渡区間を含めた評価指標の設計や、時間による動作変化をより強く捉える学習手法の導入が今後の課題である。

第二に、センサ配置および環境差に対する汎化性能の検証である。準ミリ波レーダーは、設置位置や周囲環境に依存して反射特性が変化するため、同一の動作であっても観測信号の分布が変動する可能性がある。本研究では限られた環境条件での評価に留まっているため、異なる部屋レイアウトやセンサ配置条件においてデータを収集し、モデルの性能変化を定量的に評価する必要がある。また、これらの変動に対して安定した特徴抽出を実現するためのモデル設計の検討も重要である。

第三に、少量データ条件下における学習の更なる安定化である。本研究では、ST-GCN/AGが少量データに対しても比較的安定した性能を示すことを確認したが、被験者数や動作バリエーションが増加した場合の性能については十分に検証され

ていない。特に、センサベースの姿勢推定ではデータ収集コストが高いため、限られたデータから効率的に特徴を学習する手法の重要性は高い。そのため、データ拡張や事前学習の導入、あるいはセンサ信号の構造を活用した正則化手法などを組み合わせることで、より安定した学習を実現することが求められる。

## 参考文献

- [1] Zhe Cao, Tomas Simon, Shih-En Wei, and Yaser Sheikh. Realtime multi-person 2d pose estimation using part affinity fields. In *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 1302–1310, 2017.
- [2] William McNally, Kanav Vats, Alexander Wong, and John McPhee. Rethinking keypoint representations: Modeling keypoints and poses as objects for multi-person human pose estimation. In *Computer Vision – ECCV 2022: 17th European Conference, Tel Aviv, Israel, October 23–27, 2022, Proceedings, Part VI*, page 37–54, Berlin, Heidelberg, 2022. Springer-Verlag.
- [3] Munkhjargal Gochoo, Tan-Hsu Tan, Tsedevdorj Batjargal, Oleg Seredin, and Shih-Chia Huang. Device-free non-privacy invasive indoor human posture recognition using low-resolution infrared sensor-based wireless sensor networks and dcnn. In *2018 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, pages 2311–2316, 2018.
- [4] Van Linh Nguyen, Duc Hoang Tran, Huy Nguyen, and Yeong Min Jang. Human activity detection based on infrared array sensor using advanced deep learning technique. In *2022 13th International Conference on Information and Communication Technology Convergence (ICTC)*, pages 2149–2151, 2022.
- [5] Atsuki Tanaka, Yuto Lim, and Yasuo Tan. Multi-user recognition and identification system using multiple sensors in indoor environments. In *IEICE Tech. Rep.*, volume 121 of *SeMI2021-89*, pages 35–40, Online, March 2022. Mon, Mar 7, 2022 - Tue, Mar 8 : Online (SeMI, IPSJ-MBL, IPSJ-UBI).
- [6] Saki Iwata, Yasutomo Kawanishi, Daisuke Deguchi, Ichiro Ide, Hiroshi Murase, and Tomoyoshi Aizawa. Lfir2pose: Pose estimation from an extremely low-resolution fir image sequence. In *2020 25th International Conference on Pattern Recognition (ICPR)*, pages 2597–2603, 2021.
- [7] Mingmin Zhao, Tianhong Li, Mohammad Abu Alsheikh, Yonglong Tian, Hang Zhao, Antonio Torralba, and Dina Katabi. Through-wall human pose

- estimation using radio signals. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 7356–7365, 2018.
- [8] Arindam Sengupta, Feng Jin, Renyuan Zhang, and Siyang Cao. mm-pose: Real-time human skeletal posture estimation using mmwave radars and cnns. *IEEE sensors journal*, 20(17):10032–10044, 2020.
- [9] Congzhang Ding, Shisheng Guo, Guolong Cui, and Xiaobo Yang. A parameter estimation and deep learning hybrid extraction network for multidirectional human activity recognition based on mmwave radar. *IEEE Internet of Things Journal*, 12(5):5769–5782, 2025.
- [10] Girish Tiwari and Shalabh Gupta. An mmwave radar based real-time contactless fitness tracker using deep cnns. *IEEE Sensors Journal*, 21(15):17262–17270, 2021.
- [11] Jiarui Yang, Songpengcheng Xia, Yifan Song, Qi Wu, and Ling Pei. mmbat: A multi-task framework for mmwave-based human body reconstruction and translation prediction. In *ICASSP 2024 - 2024 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 8446–8450, 2024.
- [12] David Duvenaud, Dougal Maclaurin, Jorge Aguilera-Iparraguirre, Rafael Gómez-Bombarelli, Timothy Hirzel, Alán Aspuru-Guzik, and Ryan P. Adams. Convolutional networks on graphs for learning molecular fingerprints. In *Proceedings of the 29th International Conference on Neural Information Processing Systems - Volume 2, NIPS’15*, page 2224–2232, Cambridge, MA, USA, 2015. MIT Press.
- [13] Mathias Niepert, Mohamed Ahmed, and Konstantin Kutzkov. Learning convolutional neural networks for graphs. In *Proceedings of the 33rd International Conference on International Conference on Machine Learning - Volume 48, ICML’16*, page 2014–2023. JMLR.org, 2016.
- [14] Thomas N. Kipf and Max Welling. Semi-supervised classification with graph convolutional networks, 2017.
- [15] James Atwood and Don Towsley. Diffusion-convolutional neural networks. In *Proceedings of the 30th International Conference on Neural Information Processing Systems, NIPS’16*, page 2001–2009, Red Hook, NY, USA, 2016. Curran Associates Inc.

- [16] William L. Hamilton, Rex Ying, and Jure Leskovec. Inductive representation learning on large graphs. In *Proceedings of the 31st International Conference on Neural Information Processing Systems*, NIPS’17, page 1025–1035, Red Hook, NY, USA, 2017. Curran Associates Inc.
- [17] Federico Monti, Davide Boscaini, Jonathan Masci, Emanuele Rodola, Jan Svoboda, and Michael M Bronstein. Geometric deep learning on graphs and manifolds using mixture model cnns. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 5115–5124, 2017.
- [18] Thomas Kipf, Ethan Fetaya, Kuan-Chieh Wang, Max Welling, and Richard Zemel. Neural relational inference for interacting systems, 2018.
- [19] Sijie Yan, Yuanjun Xiong, and Dahua Lin. Spatial temporal graph convolutional networks for skeleton-based action recognition. In *Proceedings of the Thirty-Second AAAI Conference on Artificial Intelligence and Thirtieth Innovative Applications of Artificial Intelligence Conference and Eighth AAAI Symposium on Educational Advances in Artificial Intelligence*, AAAI’18/IAAI’18/EAAI’18. AAAI Press, 2018.
- [20] Shaojie Bai, J Zico Kolter, and Vladlen Koltun. An empirical evaluation of generic convolutional and recurrent networks for sequence modeling. *arXiv preprint arXiv:1803.01271*, 2018.
- [21] Joan Bruna, Wojciech Zaremba, Arthur Szlam, and Yann LeCun. Spectral networks and locally connected networks on graphs, 2014.
- [22] Petar Veličković, Guillem Cucurull, Arantxa Casanova, Adriana Romero, Pietro Lio, and Yoshua Bengio. Graph attention networks. *arXiv preprint arXiv:1710.10903*, 2017.
- [23] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, and Illia Polosukhin. Attention is all you need. In I. Guyon, U. Von Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett, editors, *Advances in Neural Information Processing Systems*, volume 30. Curran Associates, Inc., 2017.
- [24] Lei Shi, Yifan Zhang, Jian Cheng, and Hanqing Lu. Two-stream adaptive graph convolutional networks for skeleton-based action recognition. In *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 12018–12027, 2019.

- [25] Shenzhen Hi-Link Electronic Co., Ltd. HLK-LD2410B+BLE 24GHz Human Presence Radar Sensor Module. <https://hlktech.net/index.php?id=1094>, n.d. Accessed: 2026-03-24.

# 謝辞

本研究を進めるにあたり, 多大なるご助言とご指導を賜りましたリム 勇仁教授, 丹 康雄教授に深く感謝申し上げます. また, 在学中に私生活の面から温かく支えてくれた家族に心より感謝いたします.

## 研究業績

1. 崎間 寛哉, 丹 康雄, リム 勇仁, ”準ミリ波レーダーを使用した姿勢推定システムの研究”, 2025 年度電気・情報関係学会北陸支部連合大会, Sep. 2025
2. 崎間 寛哉, 丹 康雄, リム 勇仁, ”準ミリ波レーダーを使用した姿勢推定システムの研究”, 電子情報通信学会 情報ネットワーク研究会, Mar. 2026

# 付録

## データ分割比率による影響

パターン A 学習データ 50%

学習過程における損失の推移

ST-GCN

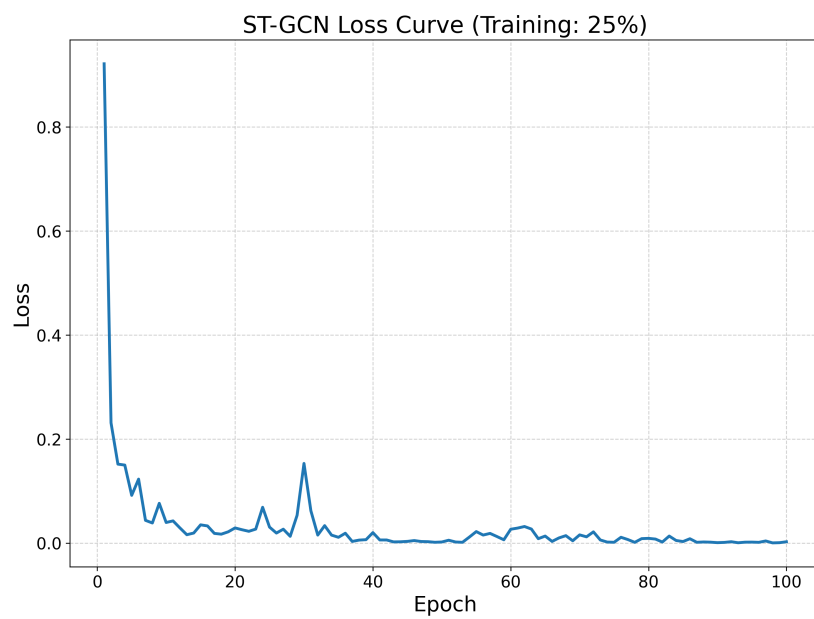


図 7.1: ST-GCN の学習における損失曲線

ST-GCN/SA

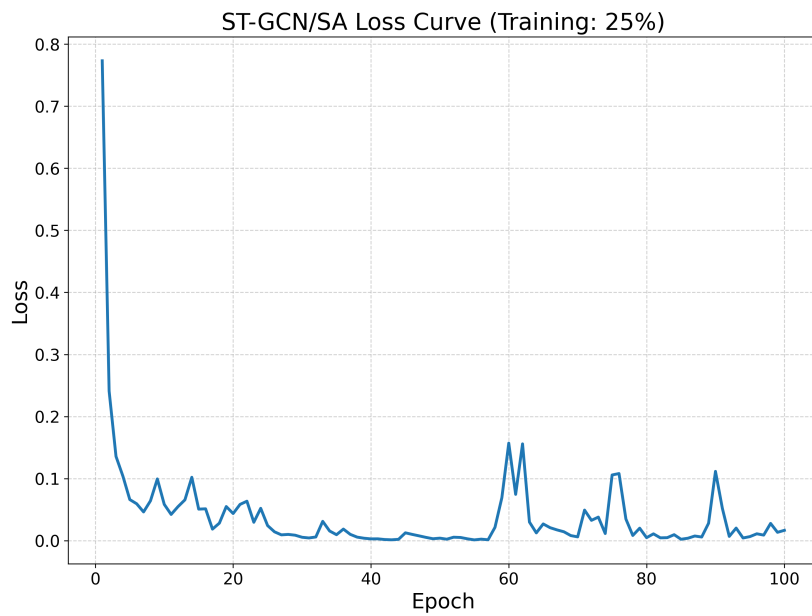


図 7.2: ST-GCN/SA の学習における損失曲線

### ST-GCN/AG

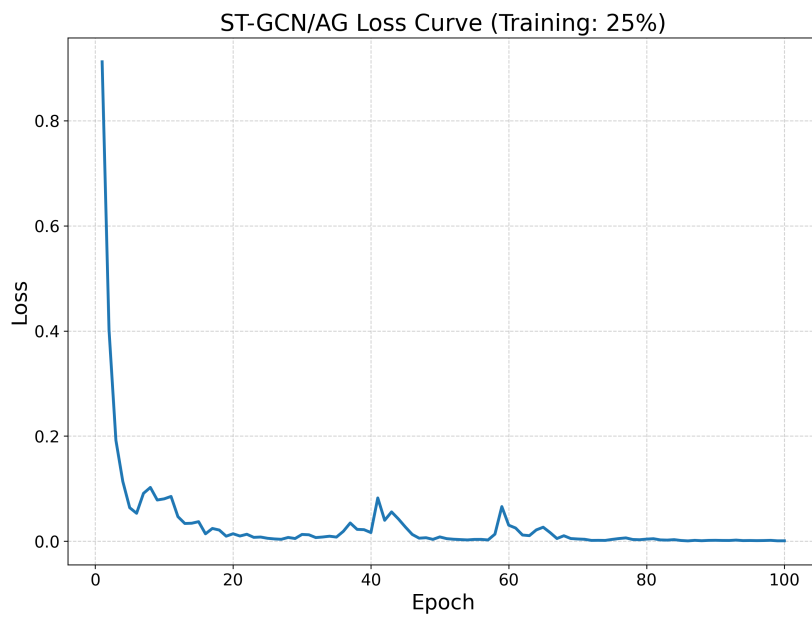


図 7.3: ST-GCN/AG の学習における損失曲線

### 単独動作時の姿勢推定の精度評価

#### ST-GCN

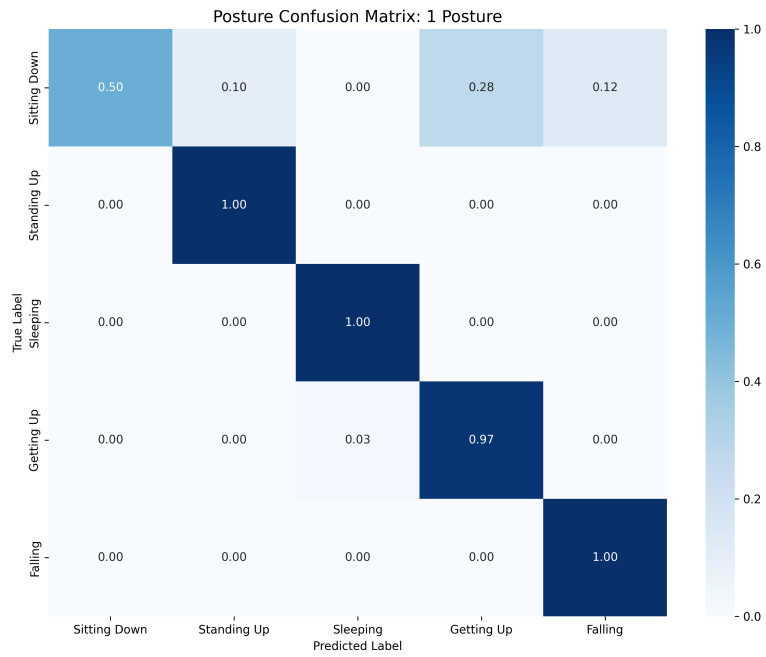


図 7.4: ST-GCN による単独姿勢動作の混同行列

### ST-GCN/SA

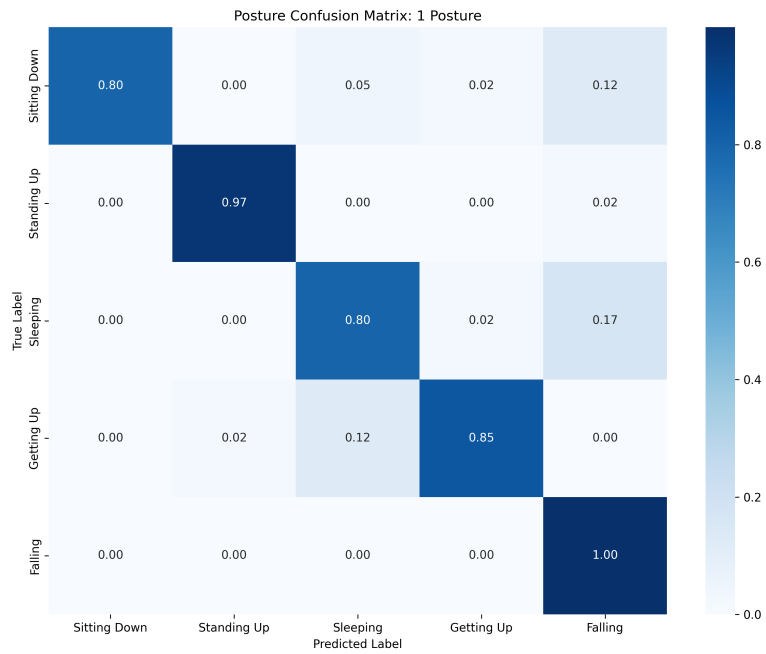


図 7.5: ST-GCN/SA による単独姿勢動作の混同行列

### ST-GCN/AG

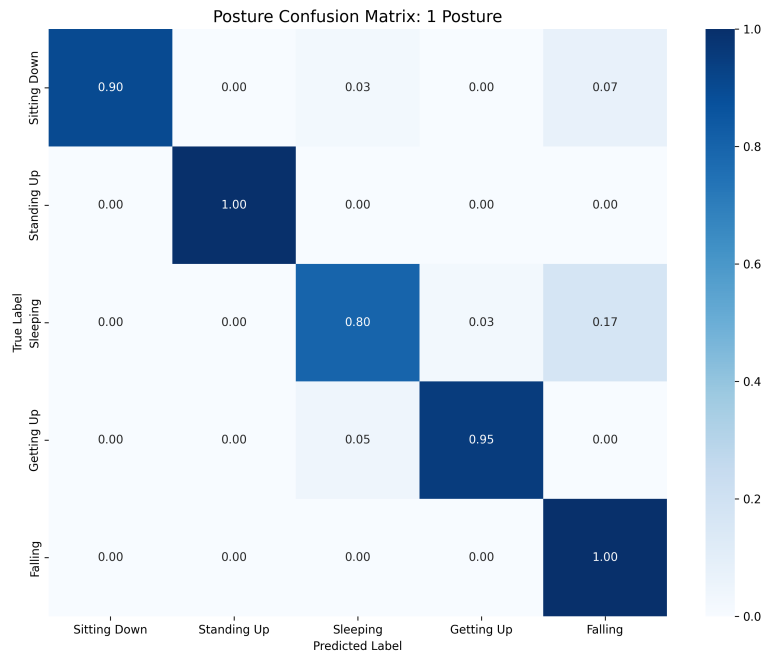


図 7.6: ST-GCN/AG による単独姿勢動作の混同行列

### 組み合わせ動作時の姿勢推定の精度評価 (2組 / 200 フレーム)

#### ST-GCN

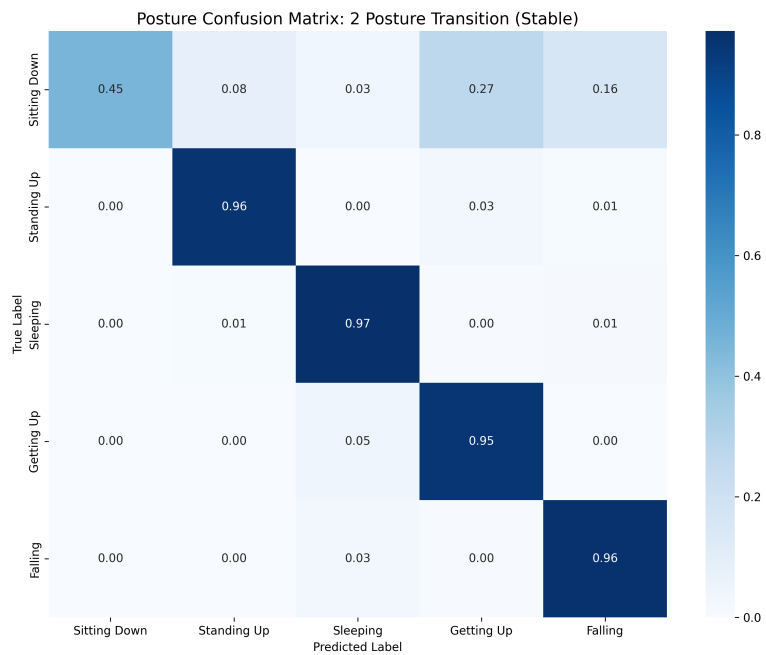


図 7.7: ST-GCN による組み合わせ姿勢動作 (2組) の混同行列

## ST-GCN/SA

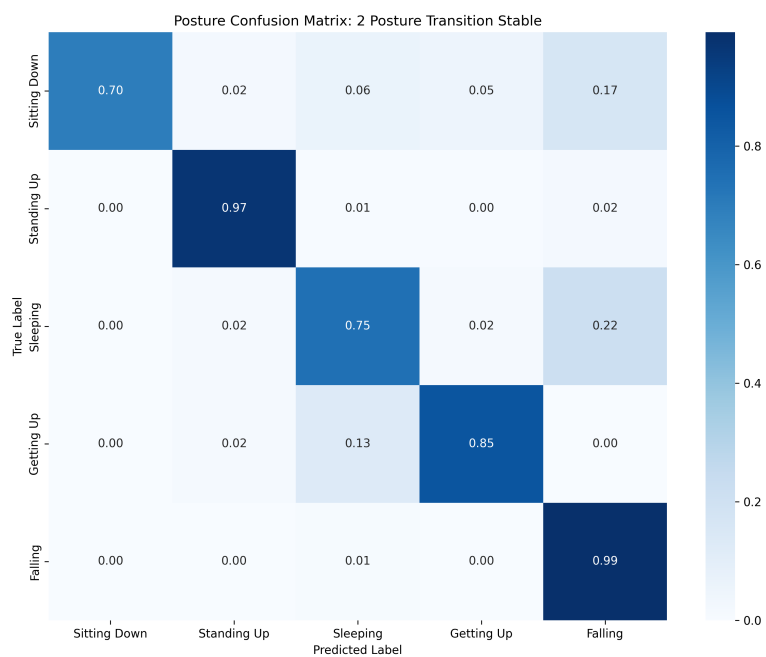


図 7.8: ST-GCN/SA による組み合わせ姿勢動作（2組）の混同行列

## ST-GCN/AG

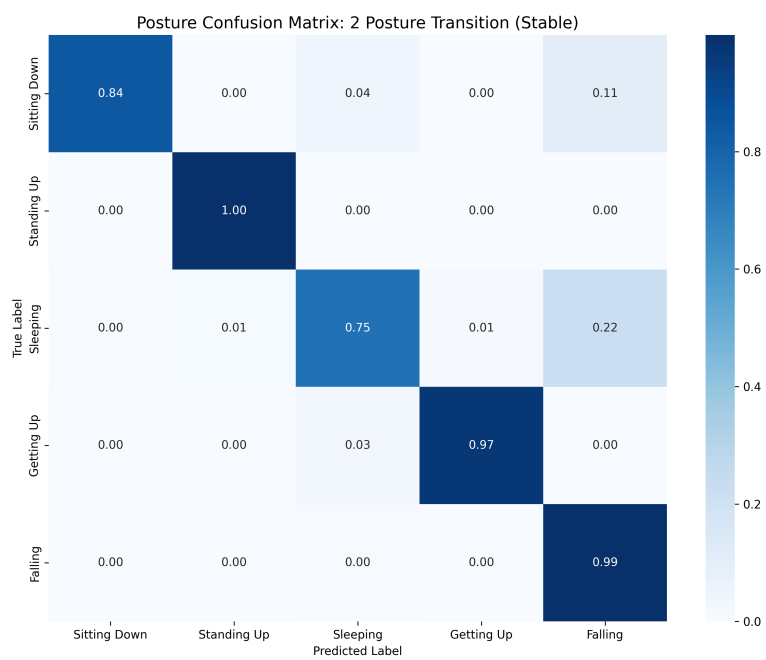


図 7.9: ST-GCN/AG による組み合わせ姿勢動作（2組）の混同行列

## 組み合わせ動作時の姿勢推定の精度評価（10組 / 1000フレーム）

### ST-GCN

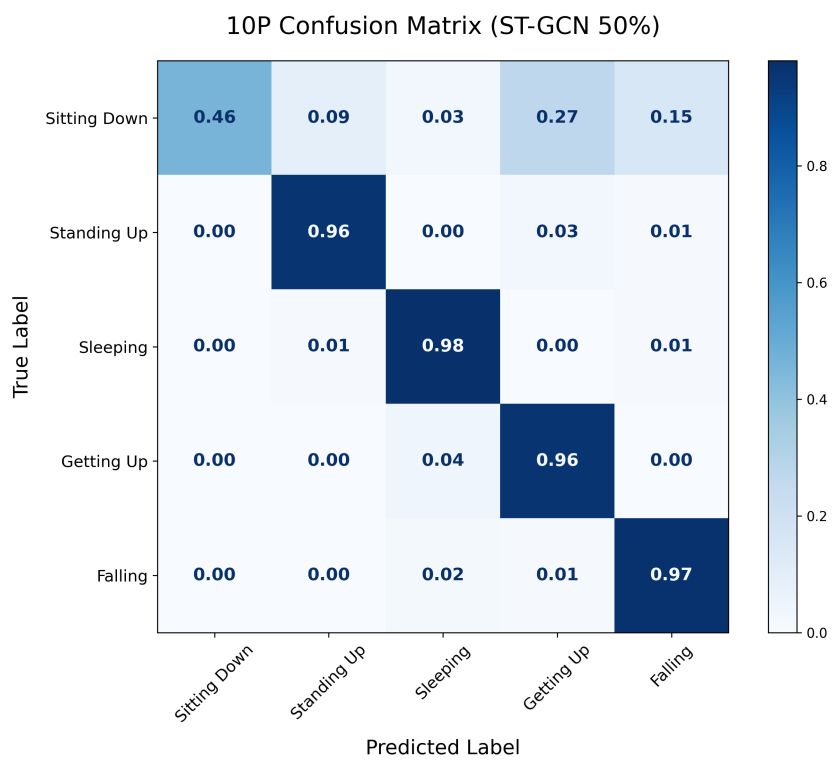


図 7.10: ST-GCN による組み合わせ姿勢動作（10組）の混同行列

### ST-GCN/SA

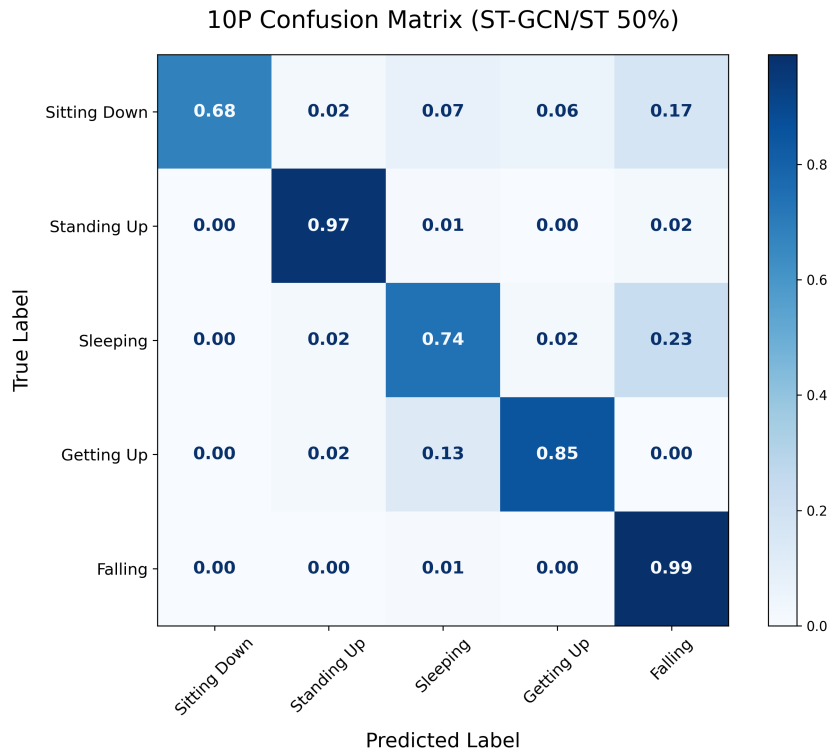


図 7.11: ST-GCN/SA による組み合わせ姿勢動作 (10 組) の混同行列

### ST-GCN/AG

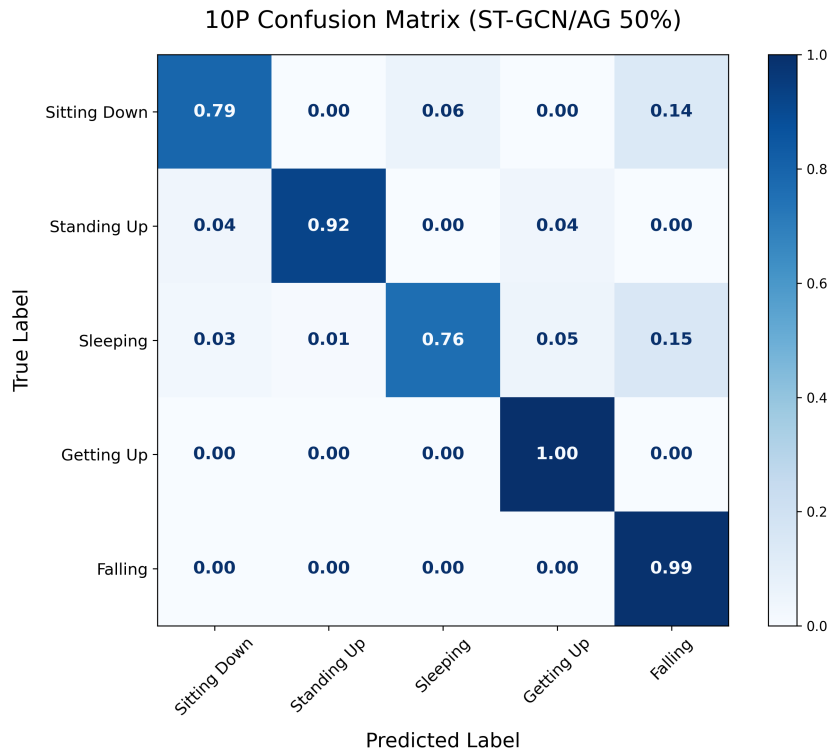


図 7.12: ST-GCN/AG による組み合わせ姿勢動作（10 組）の混同行列

## パターンB 学習データ 33%

学習過程における損失の推移

### ST-GCN

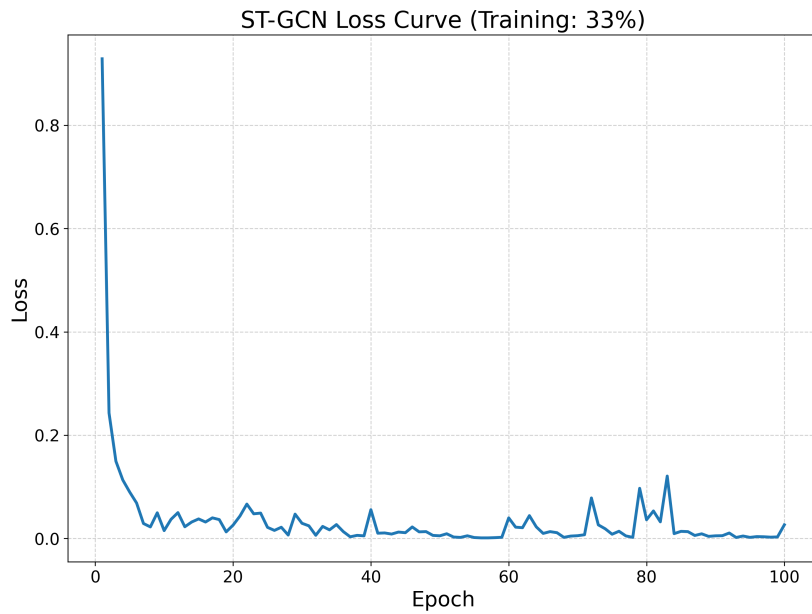


図 7.13: ST-GCN の学習における損失曲線

### ST-GCN/SA

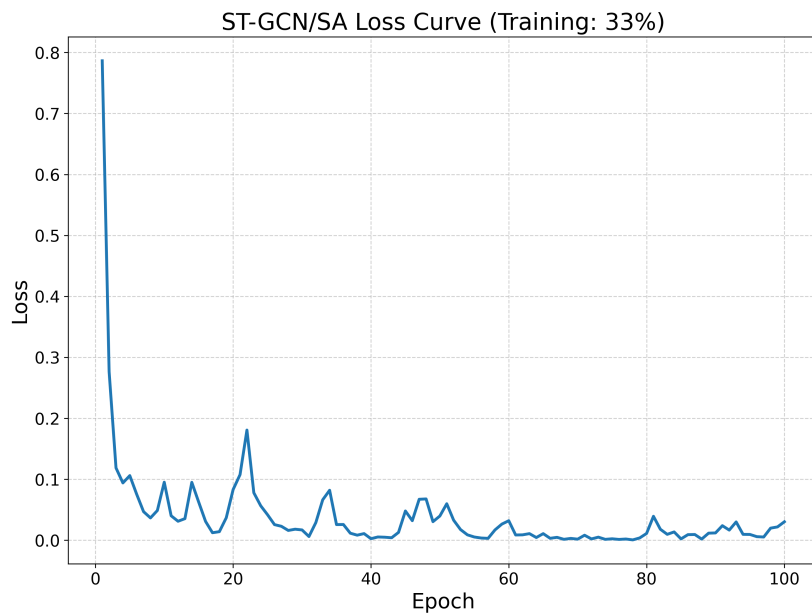


図 7.14: ST-GCN/SA の学習における損失曲線

## ST-GCN/AG

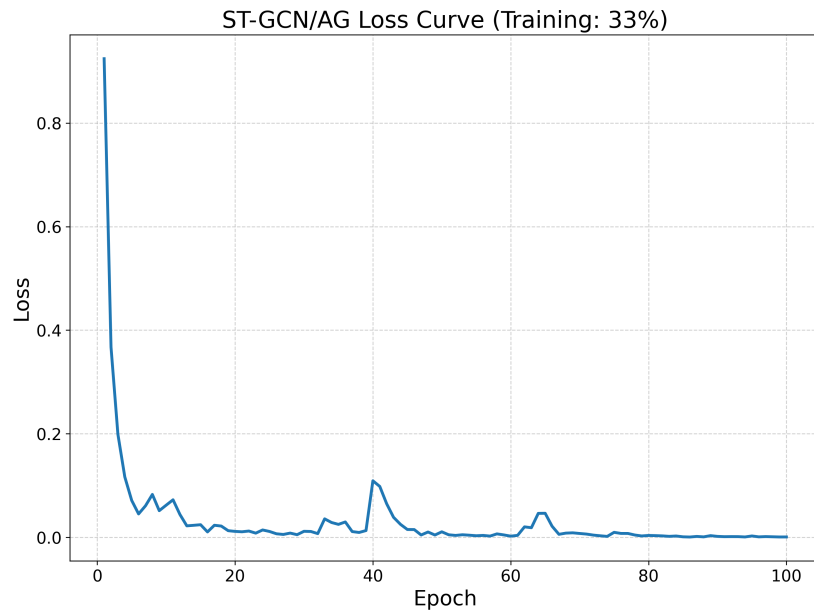


図 7.15: ST-GCN/AG の学習における損失曲線

## 単独動作時の姿勢推定の精度評価

### ST-GCN

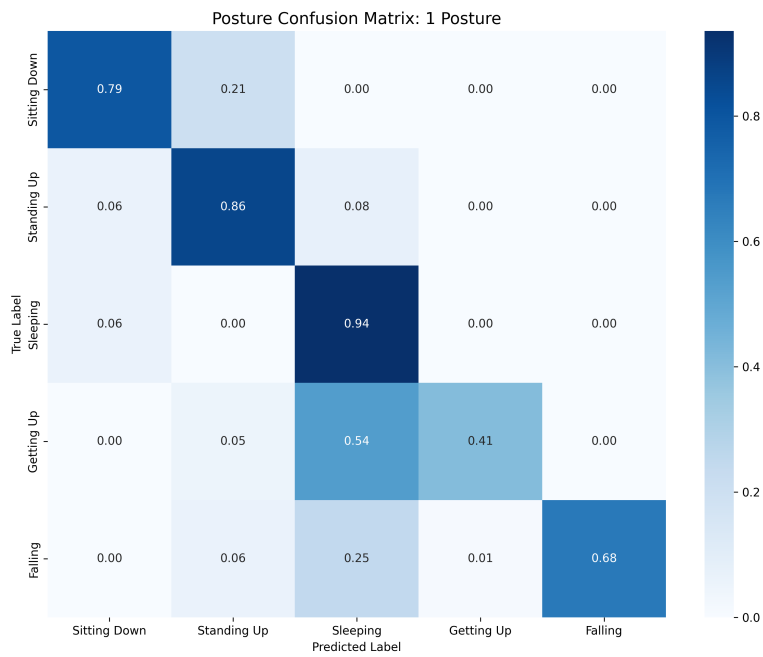


図 7.16: ST-GCN による単独姿勢動作の混同行列

## ST-GCN/SA

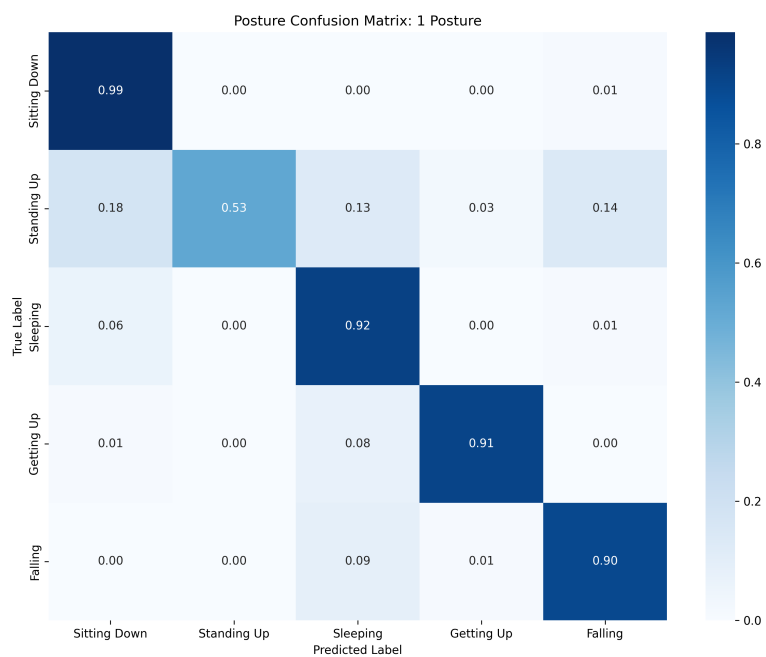


図 7.17: ST-GCN/SA による単独姿勢動作の混同行列

## ST-GCN/AG

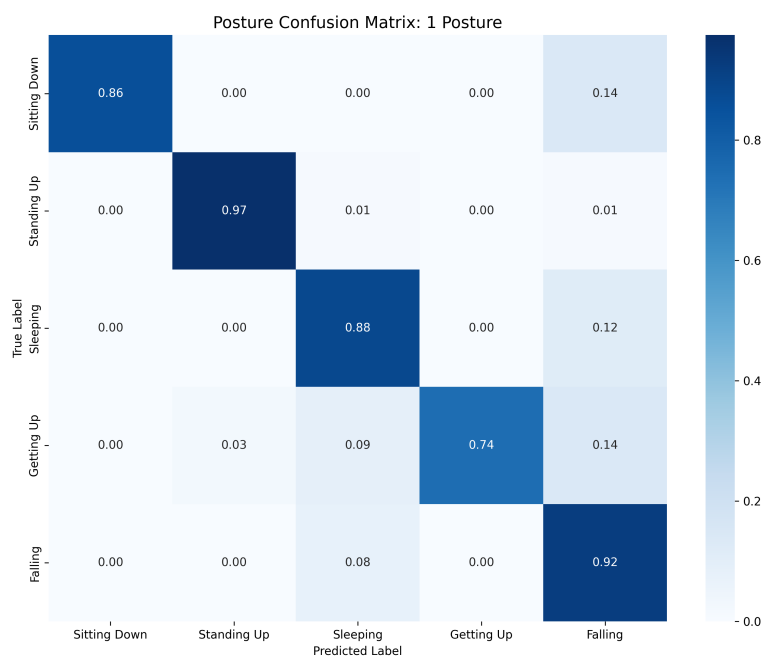


図 7.18: ST-GCN/AG による単独姿勢動作の混同行列

## 組み合わせ動作時の姿勢推定の精度評価（2組 / 200フレーム）

### ST-GCN

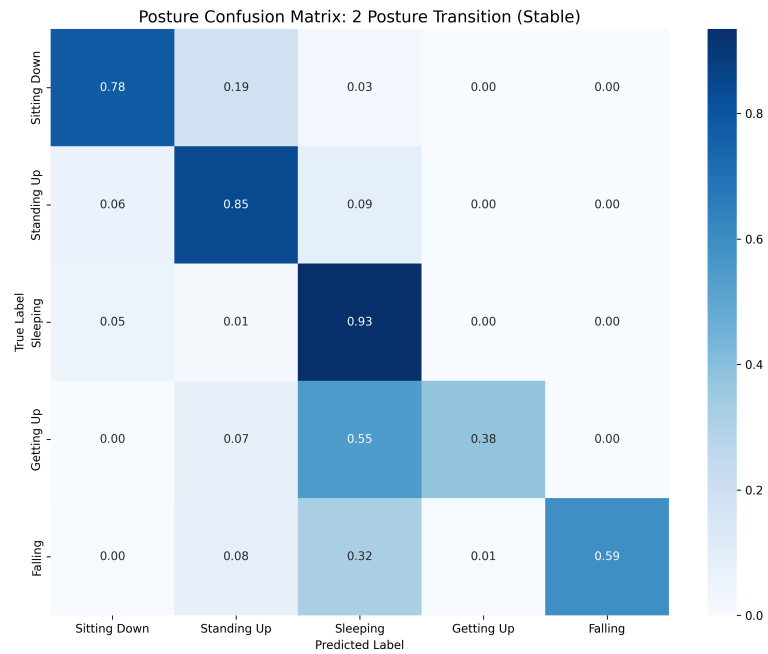


図 7.19: ST-GCN による組み合わせ姿勢動作（2組）の混同行列

### ST-GCN/SA

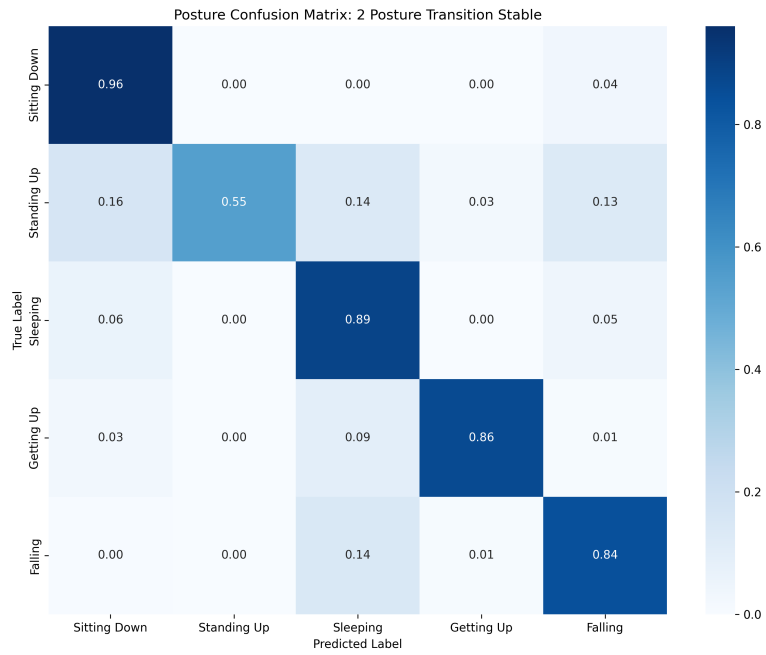


図 7.20: ST-GCN/SA による組み合わせ姿勢動作 (2 組) の混同行列

### ST-GCN/AG

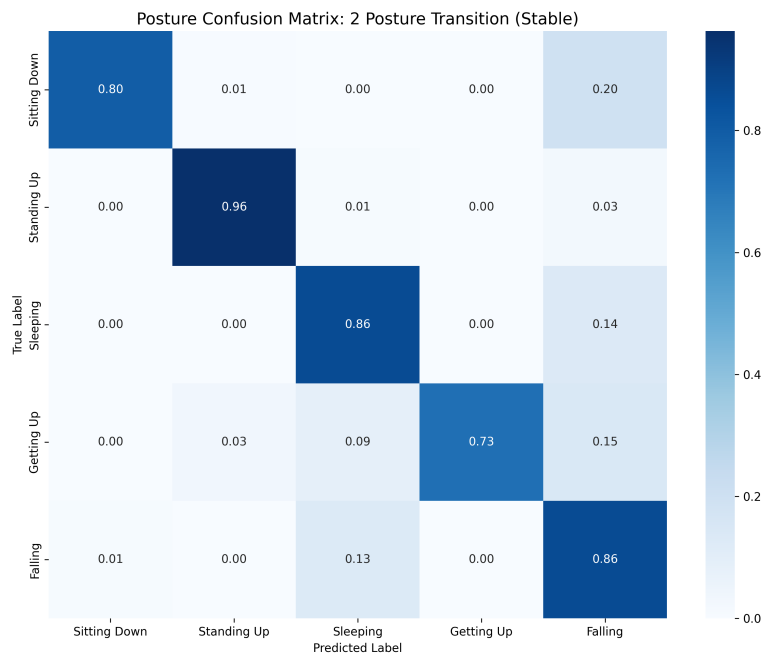


図 7.21: ST-GCN/AG による組み合わせ姿勢動作 (2 組) の混同行列

## 組み合わせ動作時の姿勢推定の精度評価（10組 / 1000フレーム）

### ST-GCN

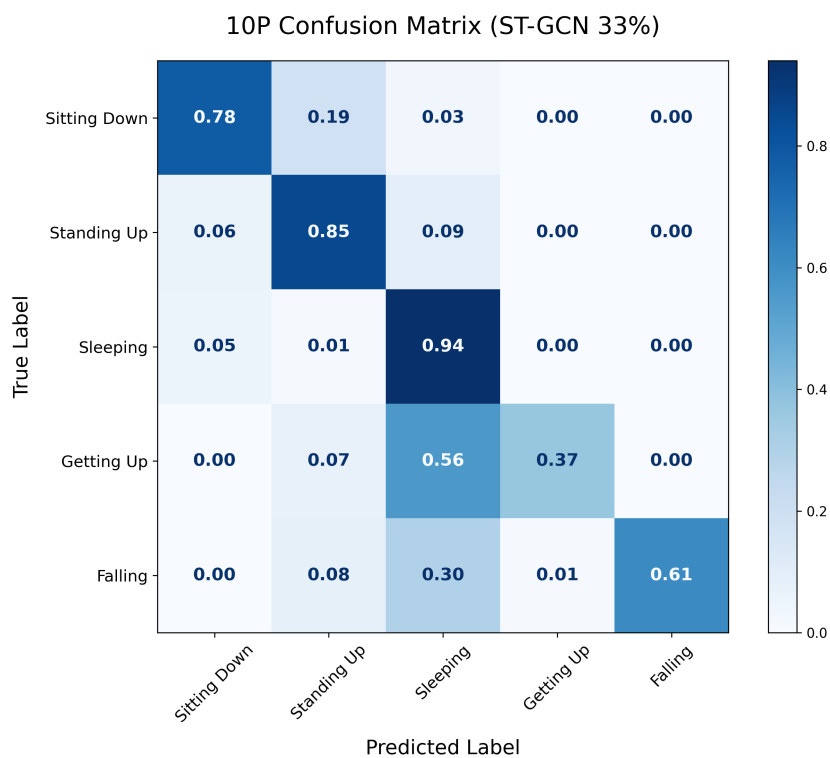


図 7.22: ST-GCN による組み合わせ姿勢動作（10組）の混同行列

### ST-GCN/SA

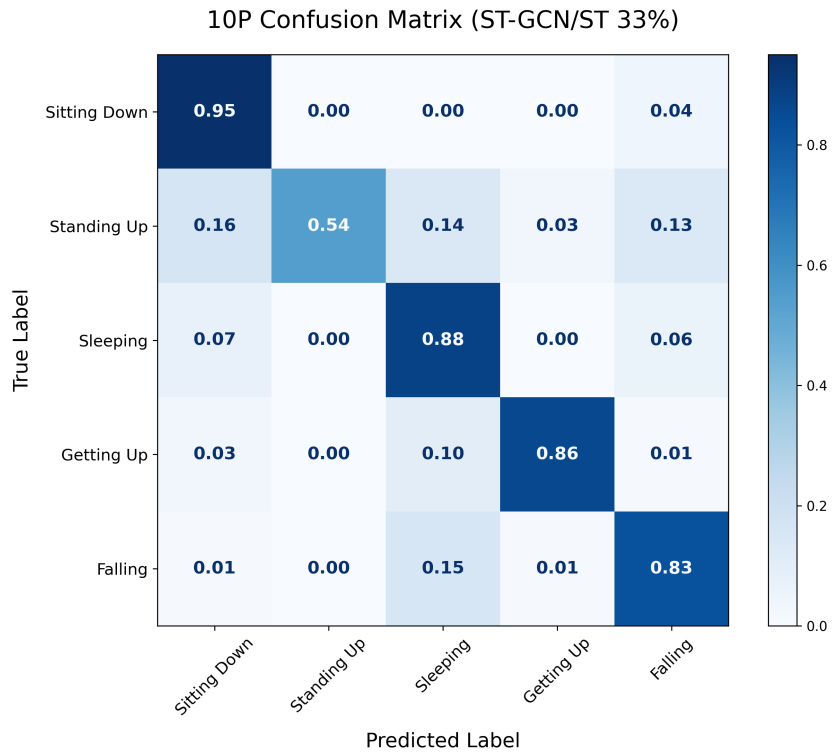


図 7.23: ST-GCN/SA による組み合わせ姿勢動作（10 組）の混同行列

### ST-GCN/AG

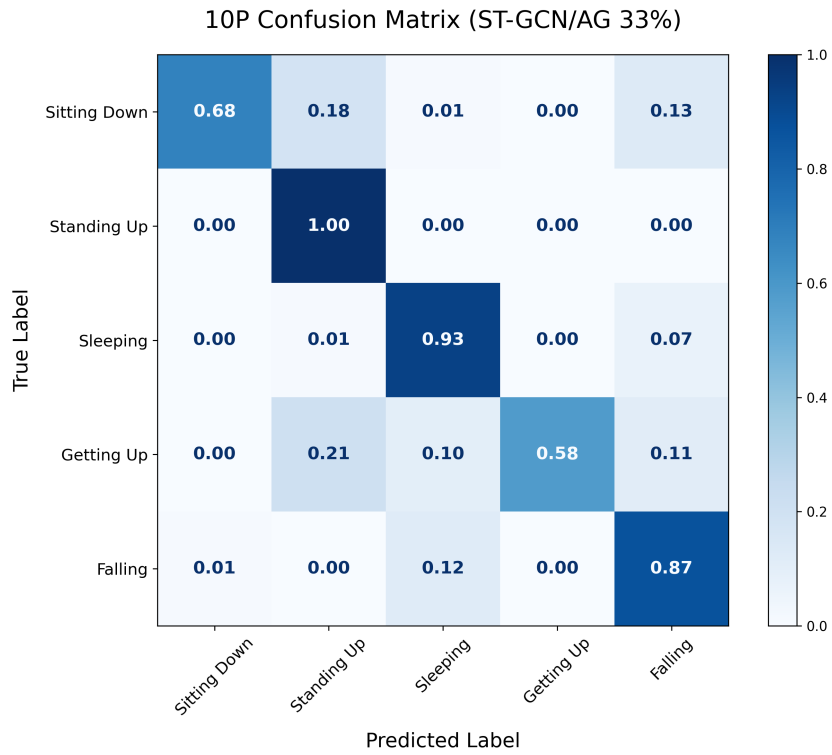


図 7.24: ST-GCN/AG による組み合わせ姿勢動作（10 組）の混同行列

## パターンC 学習データ 25%

学習過程における損失の推移

ST-GCN

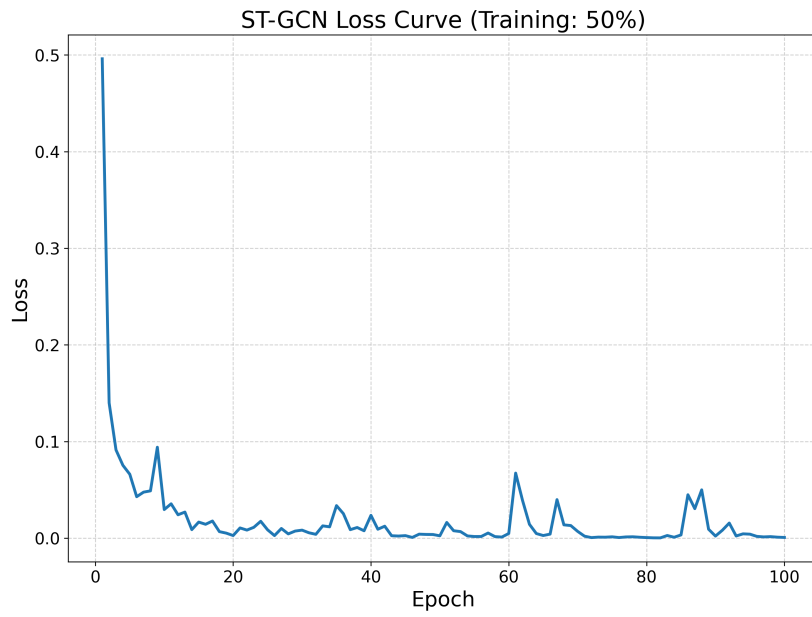


図 7.25: ST-GCN の学習における損失曲線

### ST-GCN/SA

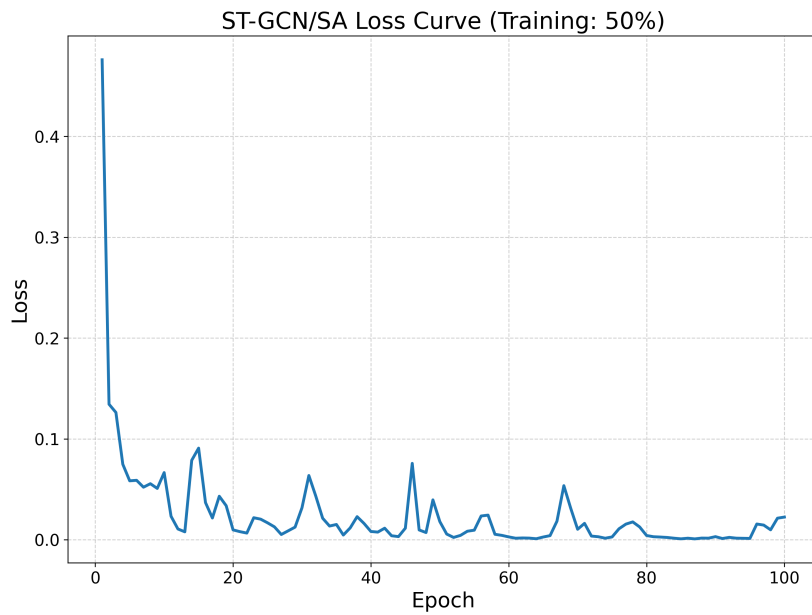


図 7.26: ST-GCN/SA の学習における損失曲線

### ST-GCN/AG

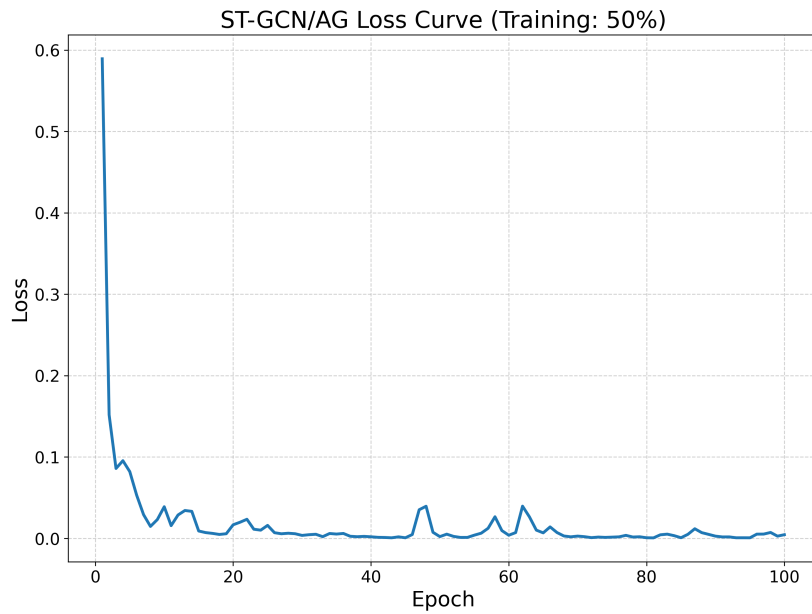


図 7.27: ST-GCN/AG の学習における損失曲線

### 単独動作時の姿勢推定の精度評価

#### ST-GCN

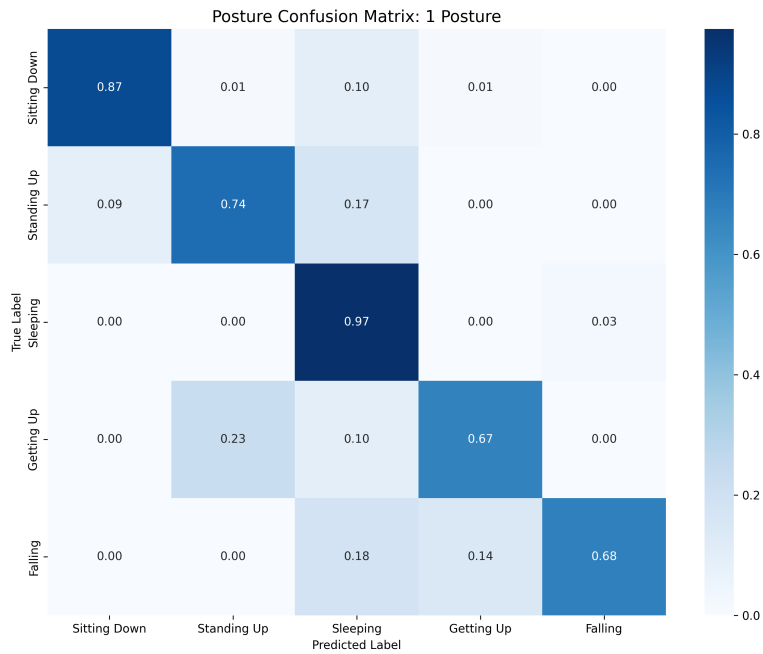


図 7.28: ST-GCN による単独姿勢動作の混同行列

## ST-GCN/SA

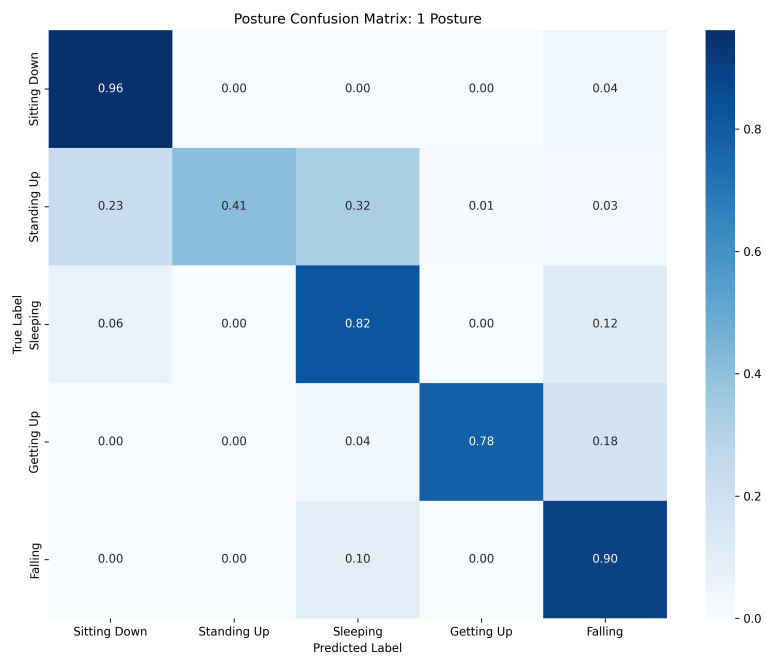


図 7.29: ST-GCN/SA による単独姿勢動作の混同行列

## ST-GCN/AG

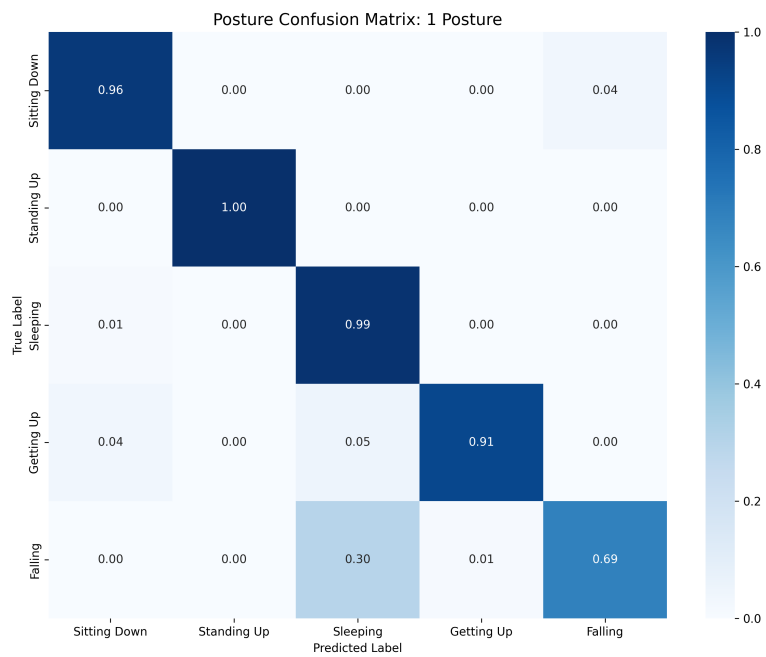


図 7.30: ST-GCN/AG による単独姿勢動作の混同行列

## 組み合わせ動作時の姿勢推定の精度評価 (2組 / 200 フレーム)

### ST-GCN

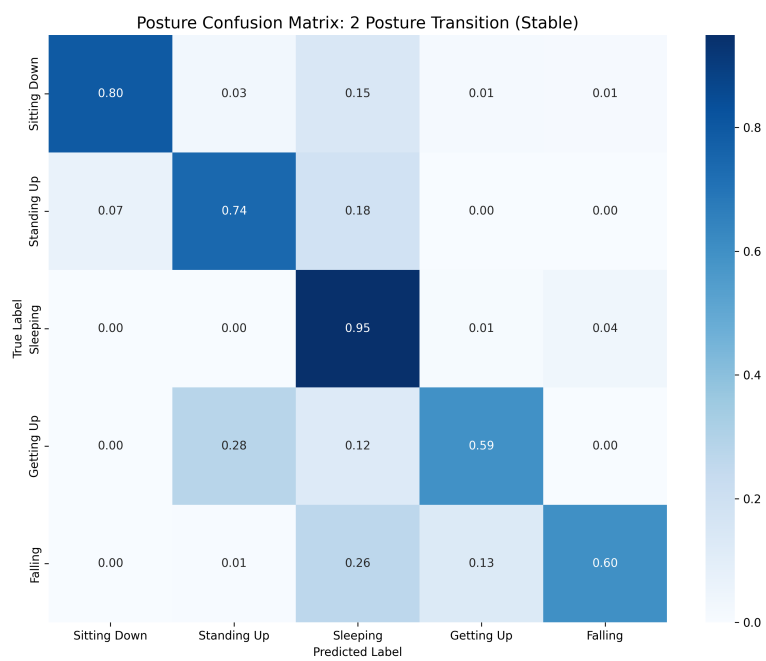


図 7.31: ST-GCN による組み合わせ姿勢動作 (2組) の混同行列

### ST-GCN/SA

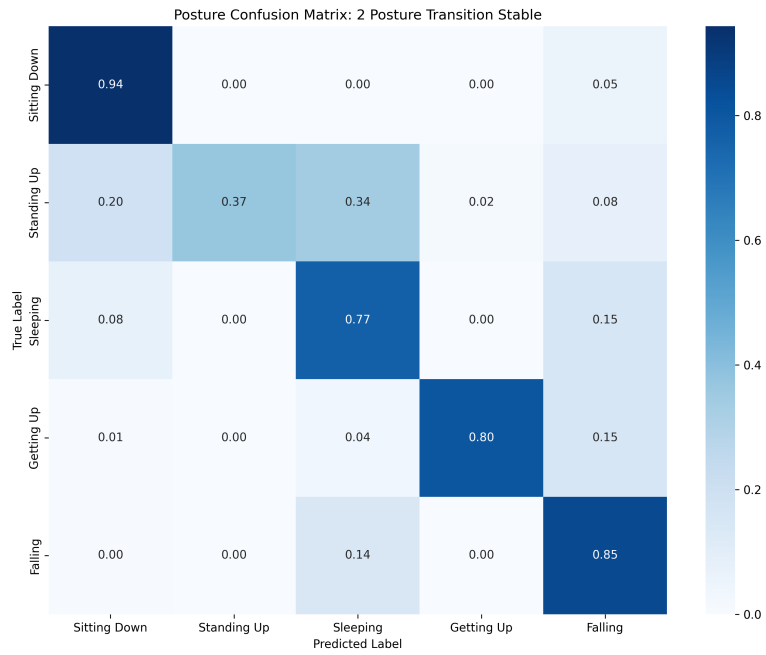


図 7.32: ST-GCN/SA による組み合わせ姿勢動作 (2 組) の混同行列

### ST-GCN/AG

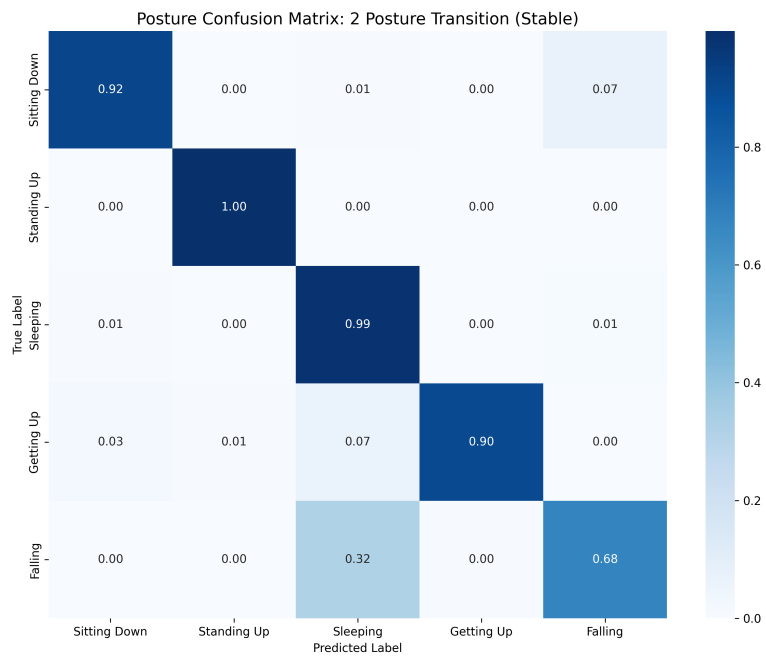


図 7.33: ST-GCN/AG による組み合わせ姿勢動作 (2 組) の混同行列

## 組み合わせ動作時の姿勢推定の精度評価（10組 / 1000フレーム）

### ST-GCN

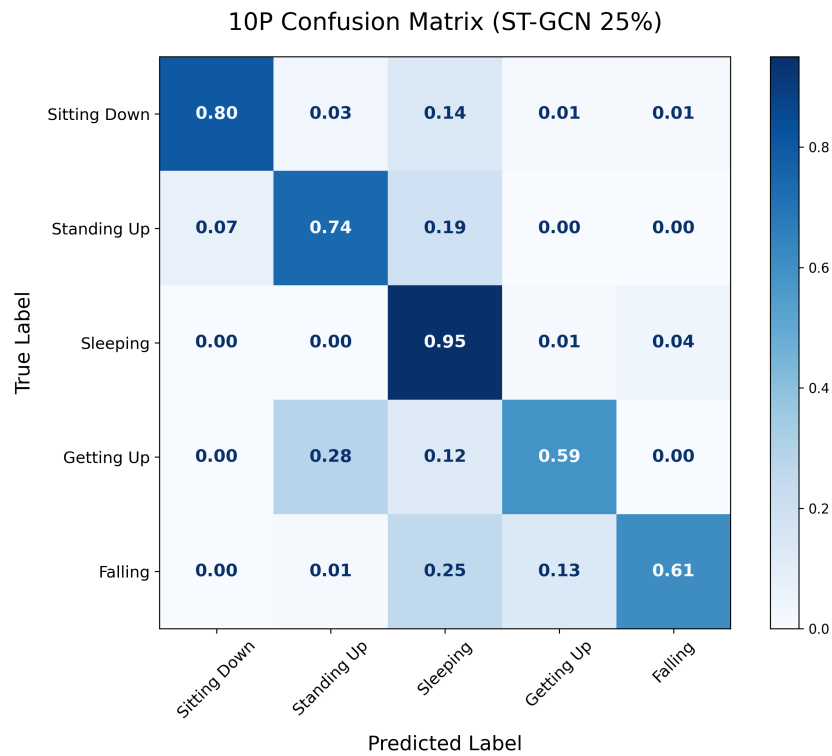


図 7.34: ST-GCN による組み合わせ姿勢動作（10組）の混同行列

### ST-GCN/SA

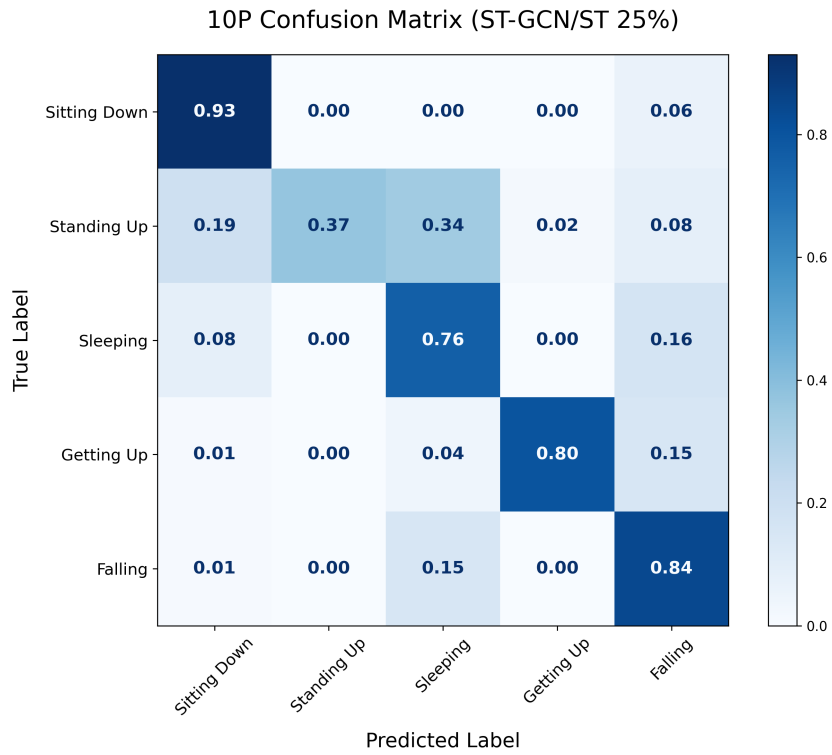


図 7.35: ST-GCN/SA による組み合わせ姿勢動作（10 組）の混同行列

### ST-GCN/AG

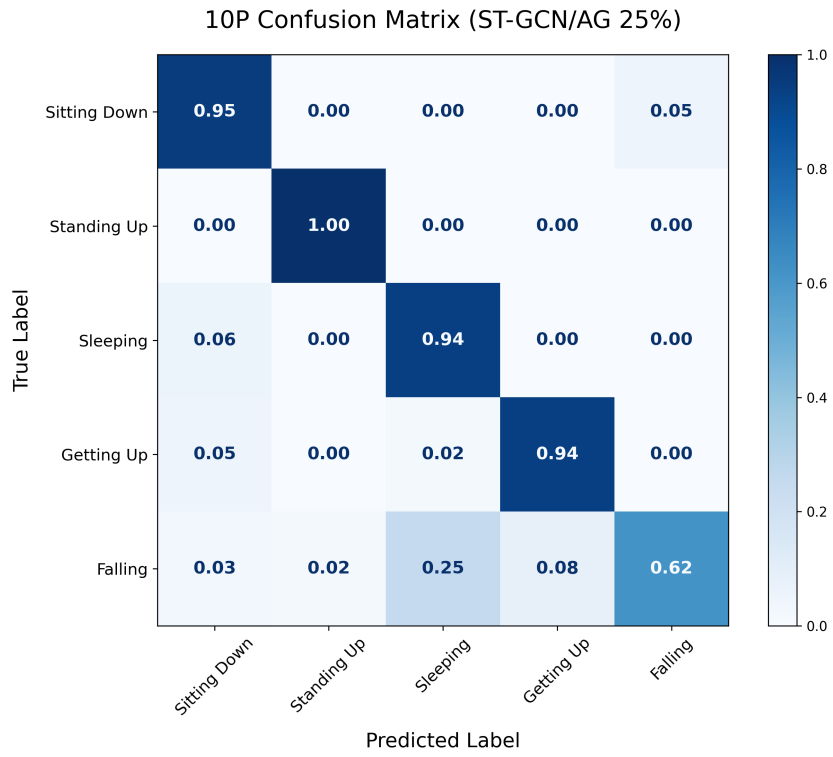


図 7.36: ST-GCN/AG による組み合わせ姿勢動作（10 組）の混同行列