

介護における音響 HAR と連合学習を用いた異常検知

発表者名（竹本志恩様の例に倣い、発表者名はプレースホルダーとしました）

June 4, 2025

所属機関名（INIAD 様の例に倣い、所属機関名はプレースホルダーとしました）

はじめに：背景と目的

関連研究と背景

技術要素と手法

課題

将来展望

結論

はじめに：背景と目的

- 増大する高齢者ケアの必要性
 - 世界的な高齢化の進行に伴い、遠隔・継続的な健康モニタリングの需要が増大。
 - 特に単身高齢者の安全確保とタイムリーな介入が課題。
- 従来のモニタリング手法の課題
 - プライバシーへの配慮や、重要な事象を捉える上での限界がある。
 - カメラベースのシステムはプライバシー侵害の懸念が大きい。
- 音響 HAR の可能性
 - 非侵入的なモニタリング手段として有望視されている。
 - 環境音から活動、安全、健康状態に関する豊富な情報を得られる。
- プライバシー保護 AI としての連合学習 (FL)
 - 生の機密性の高いデータを共有せずにモデルを訓練する手法。
 - 分散データソース（各家庭やデバイス）で協調学習を行う。
 - 音響データのような機密情報に適している。

目次

はじめに：背景と目的

関連研究と背景

技術要素と手法

課題

将来展望

結論

- 人間行動認識 (HAR) の概要
 - 生センサー入力から人間の活動に関する高レベルの知識を学習する。
 - ウェアラブル、環境センサー、ビデオなど多様なモダリティが利用される。
- 従来の HAR における課題
 - ヒューリスティックな手動特徴抽出に依存し、ドメイン知識が必要な場合がある。
 - 浅い特徴しか学習できない場合がある。
- 連合学習 (FL) の基本
 - 分散されたクライアントデータで、モデル更新の集約を通じてグローバルモデルを学習。
 - データの集中収集に伴う多くのプライバシーリスクとコストを軽減。
 - Federated Averaging (FedAvg) は代表的なアルゴリズム。
- HAR における FL の必要性
 - HAR データはユーザーや環境によって非 IID (統計的に不均一) かつ不均

目次

はじめに：背景と目的

関連研究と背景

技術要素と手法

課題

将来展望

結論

- 介護における重要な音響イベント
 - 緊急事態の直接的指標: 転倒音、悲鳴・叫び声、助けを求める言葉、異常な呼吸音。
 - 苦痛・不快感: うめき声・呻き声、泣き声、持続的な咳。
 - 危険を示唆する環境音: ガラス破損音、煙/火災警報、長時間放置された流水音。
 - 日常活動からの逸脱: 異常な時間帯の音、通常と異なるパターンの音。
- 音響特徴抽出
 - プライバシー保護のため、オンデバイスでの抽出が不可欠。
 - メル周波数ケプストラム係数 (MFCCs): 音声/音響処理で一般的、オンデバイスに適応。
 - メルスペクトログラム: CNN 入力に適し一般的だが、MFCC より情報漏洩の懸念があるとの指摘も。
- FL アルゴリズム

目次

はじめに：背景と目的

関連研究と背景

技術要素と手法

課題

将来展望

結論

- データに関する課題
 - 実世界の緊急イベントデータの希少性: シミュレーション/実験データへの依存が高い。
 - アノテーションの質と粒度: 複雑な行動や連続的なイベントのアノテーションが困難。
 - データ不均衡: 緊急イベントは稀なため、データセットが偏る。
 - データ異質性 (Non-IID) とドメインシフト: 環境音響特性やユーザー行動の多様性が課題。
 - データ不足: FL のための大規模で多様な実環境データセットが不足。
- 技術的な課題
 - ノイズ耐性: 実環境はノイズが多く、関連音の識別が必要。
 - 計算・通信リソース: エッジデバイスの制限、モデル更新の通信コスト。ローカル計算量の増加は通信効率化に寄与するが、デバイス性能とのバランスが必要。

目次

はじめに：背景と目的

関連研究と背景

技術要素と手法

課題

将来展望

結論

- データセットの拡充
 - 実世界の緊急事態を反映した、多様で高品質なアノテーション付きデータセットの必要性。
 - 文脈情報を重視したアノテーション手法の開発。
 - 合成データ生成技術の活用。
- FL アルゴリズムと PETs の進化
 - 音響データに特化した効率的な PETs の研究。
 - プライバシーとモデル有用性のバランス最適化。
 - 非 IID データやデータ不均衡に対する頑健な FL/pFL 手法の開発。
 - リソース制約のあるデバイス向け FL の最適化。
- 透明性・説明可能性 (XAI)
 - 音響 SED モデルのための XAI 技術の研究・統合。
 - 解釈可能なモデルの選択。
- マルチモーダル融合

目次

はじめに：背景と目的

関連研究と背景

技術要素と手法

課題

将来展望

結論

- 音響 HAR と FL は、高齢者介護におけるプライバシー保護異常検知システムの構築において大きな可能性を持つ。
- 転倒、苦痛の音声、異常呼吸音など緊急性の高い音響イベントの正確な認識が重要。
- 成功には、適切な音響特徴抽出、FL アルゴリズム、そして PETs の適用が不可欠だが、トレードオフの考慮が必要。
- 実世界データの不足、非 IID 性、リソース制約、倫理的課題など、克服すべき重要な課題が多数存在する。
- 今後の研究は、データ拡充、FL/PETs の最適化、XAI、マルチモーダル融合、そして倫理的配慮を組み込んだ実環境展開戦略に焦点を当てる必要がある。
- これらの課題への取り組みを通じて、FL と音響 HAR は高齢者の安全と QOL 向上に貢献できると期待される。