CareAttenNet: Deep Learning Framework with Temporal Attention for Automated Nursing Activity Recognition from Wearable Sensors

1st Xiaoyan Li†
Department of Automation
Tsinghua University
Beijing, China
xiaoyanli629@tsinghua.edu.cn

2nd Cuicui Jiang†
College of Computer Science
Inner Mongolia University
Hohhot, China
jiancuicui@mail.imu.edu.cn

3rd Zhiming Jiang
Technical Innovation Dept.
Pacvue
Wuxi, China
daniel.jiang@pacvue.com

4th Caihong Li
Department of Nursing
Tsinghua University
Beijing, China
lcha00318@btch.edu.cn

5th Zhenzhen Xie
School of Computer Science
Shandong University
Qingdao, China
xiezz21@sdu.edu.cn

6th Li Liu
Department of Automation
Tsinghua University
Beijing, China
liuli95@mail.tsinghua.edu.cn

7th Chuan Song
Institute of Flexible Electronics
Northwestern Polytechnical Univ.
Xi'an, China
chuansong@mail.nwpu.edu.cn

8th Zijing Chen*
School of Clinical Medicine
Tsinghua University
Beijing, China
rlahr@a2gov.org

2025 IEEE International Conference on Bioinformatics and Biomedicine (BIBM) | 979-8-3315-1557-7/25/$31.00 ©2025 IEEE | DOI: 10.1109/BIBM66473.2025.11356292

***Abstract*—Traditional nursing activity recognition relies on manual observation and documentation, which are time-consuming and error-prone. This paper presents CareAttenNet, a deep learning framework integrating adaptive feature selection, correlation-aware processing, and temporal attention mechanisms for automated nursing activity recognition from wearable sensor data. We evaluated the framework using the SONAR dataset comprising 70-dimensional sensor features from 14 healthcare professionals performing 23 nursing activities, totaling 7,631,843 temporal measurements. CareAttenNet achieved 77.36% validation accuracy and 60.00% test accuracy, outperforming baseline architectures including CNN-LSTM (57.92%), Correlation-Aware CNN (54.62%), and Feature-Selective Network (52.86%). Ablation studies revealed temporal attention as the most effective component (78.33% test accuracy), while feature selection combined with temporal attention achieved 77.40% accuracy. However, combining all components resulted in performance degradation, indicating complex negative interactions between architectural modules. These findings provide insights into multi-modal sensor fusion challenges and establish a foundation for intelligent healthcare monitoring systems.**

***Index Terms*—Nursing activity recognition, Wearable sensors, Deep learning, Temporal attention, Healthcare monitoring**

I. Introduction

Healthcare systems worldwide face unprecedented challenges in maintaining quality care while managing increasing patient loads and resource constraints [1, 2]. Nursing activities constitute a critical component of healthcare delivery, encompassing complex sequences of patient care tasks that require precise execution and continuous monitoring [3]. Traditional approaches to nursing activity recognition rely heavily on manual observation and documentation, which are inherently subjective, time-consuming, and prone to human error [4, 5]. The emergence of wearable sensor technologies and advanced machine learning techniques presents unprecedented opportunities to revolutionize nursing care through automated activity recognition systems [6, 7].

The automatic recognition of nursing activities from sensor data faces several critical challenges. Nursing activities involve complex temporal patterns and subtle variations that complicate classification [8]. Healthcare sensor data presents high dimensionality and noise, with significant inter-subject variability [9]. Additionally, conventional machine learning approaches often overlook the physical and semantic constraints in human motion data [10]. Human activity recognition has been extensively studied for daily and sports activities [11], however, existing approaches to nursing activity recognition remain limited by single-technique methods [12, 13] and methodological issues including data leakage and inadequate evaluation protocols [14, 15]. This study presents CareAttenNet, a comprehensive deep learning framework for nursing activity recognition that addresses existing challenges through innovative architectural design. The framework builds upon and extends established neural architectures including adaptive feature selection, correlation-aware processing and temporal attention modeling. The key contributions include novel architectural innovations incorporating domain-specific sensor physics knowledge through physics-informed feature grouping [16] and adaptive attention mechanisms [17]. The framework is evaluated through rigorous experimental protocols with comprehensive regularization strategies [18], while systematic ablation studies analyze the effectiveness of each component [19]. Our evaluation utilizes the SONAR (Sensor-Oriented Nursing Activity Recognition) dataset [20] which represents one of the most comprehensive collections of real-world nursing activity data. The data collection method is

† Equal contribution.
* Corresponding Author: Zijing Chen(zijingc@tsinghua.edu.cn)

shown in Figure 2A. The dataset comprises 70-dimensional sensor features (Figure 2B) captured from 14 healthcare professionals performing 23 distinct nursing activities in authentic clinical environments.

II. Related Work

A. Patient Activity Analysis in Clinical Nursing Research

Nursing research has explored various approaches to document and analyze patient activities, with observational studies capturing patient behaviors through comprehensive activity logs and standardized assessment tools that focus on key activities of daily living and physical movements [21, 22]. Established instruments such as the Barthel Index and Functional Independence Measure (FIM) have become widely adopted for assessing patient activity levels and functional status in clinical settings [23, 24].

Clinical studies have demonstrated the application of activity analysis in specific patient populations. Research in post-stroke rehabilitation has shown that systematic recording of patient movements helps evaluate recovery progress and guide therapeutic interventions [25]. In geriatric care, activity pattern analysis has been used to assess fall risks and mobility status, contributing to preventive care strategies [26, 27].

Activity documentation and analysis have influenced evidence-based nursing practice. Studies have shown that structured observation of patient behaviors can improve the timing and effectiveness of nursing interventions [28, 29].

B. Applications of AI in Clinical Nursing Practice

Artificial intelligence [30] has advanced patient activity monitoring in nursing practice through machine learning algorithms [31] and neural networks [32]. However, existing applications often focus on single-task recognition without considering the sequential nature of nursing activities [33, 34].

Deep learning approaches have demonstrated effectiveness in analyzing patient behavior patterns and predicting health events [35, 36]. While transformers have enhanced temporal pattern analysis [37], most approaches overlook nursing-specific characteristics and clinical constraints [38].

AI systems support clinical decision-making by analyzing patient activity patterns [39, 40]. However, existing solutions often use simplified models that cannot fully capture real-world nursing complexity [41] or handle multi-modal data streams with hierarchical task structures [42].

The combination of AI techniques with nursing expertise has led to improved patient monitoring systems. These approaches have demonstrated effectiveness in detecting and classifying patient activities, supporting evidence-based nursing interventions [43]. However, previous work has typically employed generic activity recognition models that do not incorporate nursing-specific knowledge or workflow patterns [44]. The integration of domain expertise with advanced AI architectures remains an underexplored area that could significantly enhance the accuracy and reliability of activity recognition systems in nursing applications [45].

III. Methods

A. Problem Formulation

The nursing activity recognition problem is formalized as a multivariate time series classification task. At each time step t, an inertial measurement unit (IMU) system generates a D-dimensional feature vector $\mathbf{x}_t \in \mathbb{R}^D$, where the dimensionality $D = 70$ corresponds to multi-modal features obtained from 5 body-worn sensors.

Problem Definition: Given a temporal observation window of length T, denoted as $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_T] \in \mathbb{R}^{T\times D}$, we seek to learn a mapping function:

$$f : \mathbb{R}^{T\times D} \to \mathbb{R}^C \tag{1}$$

that maps the temporal window to a probability distribution over C nursing activity categories, where $C = 23$ represents distinct nursing activity types.

Multi-modal Sensor Structure: The sensor data exhibits a clear physical structure reflecting different aspects of human motion:

$$\mathbf{x}_t = \begin{bmatrix} \mathbf{q}_t^{(1)}, \mathbf{q}_t^{(2)}, \ldots, \mathbf{q}_t^{(S)} \\ \dot{\mathbf{q}}_t^{(1)}, \dot{\mathbf{q}}_t^{(2)}, \ldots, \dot{\mathbf{q}}_t^{(S)} \\ \mathbf{v}_t^{(1)}, \mathbf{v}_t^{(2)}, \ldots, \mathbf{v}_t^{(S)} \\ \mathbf{m}_t^{(1)}, \mathbf{m}_t^{(2)}, \ldots, \mathbf{m}_t^{(S)} \end{bmatrix} \tag{2}$$

where $S = 5$ represents the number of sensors, and for each sensor s: $\mathbf{q}_t^{(s)} \in \mathbb{R}^4$ denotes quaternion orientation representation, $\dot{\mathbf{q}}_t^{(s)} \in \mathbb{R}^4$ represents quaternion derivatives (angular velocity-related), $\mathbf{v}_t^{(s)} \in \mathbb{R}^3$ captures linear velocity and acceleration components, and $\mathbf{m}_t^{(s)} \in \mathbb{R}^3$ measures three-axis magnetic field intensity.

B. Dataset

This study utilized the SONAR dataset, which represents one of the most comprehensive collections of real-world nursing activity data captured in authentic clinical environments. The dataset was collected in a single retirement and assisted living facility in Potsdam, Germany, where experienced professional nurses (aged 24-59, representing diverse genders) provided care to elderly residents with physical or mental limitations who resided in single rooms and required assistance with daily living activities. The dataset comprises sensor measurements collected from 14 healthcare professionals performing 23 distinct nursing activities using wearable inertial measurement units (IMUs). This dataset contains a total of 7,631,843 temporal measurements across all subjects and activities, with each measurement consisting of 70-dimensional sensor feature vectors. These features are organized into four physically meaningful groups: quaternion orientation data (12 dimensions representing device orientation in 3D space), quaternion derivatives (12 dimensions capturing orientation change rates), velocity measurements (24 dimensions including linear and angular velocities), and magnetic field data (23 dimensions from magnetometer readings). The nursing activities were labeled in real-time by an external observer who walked alongside the nurses and recorded activity

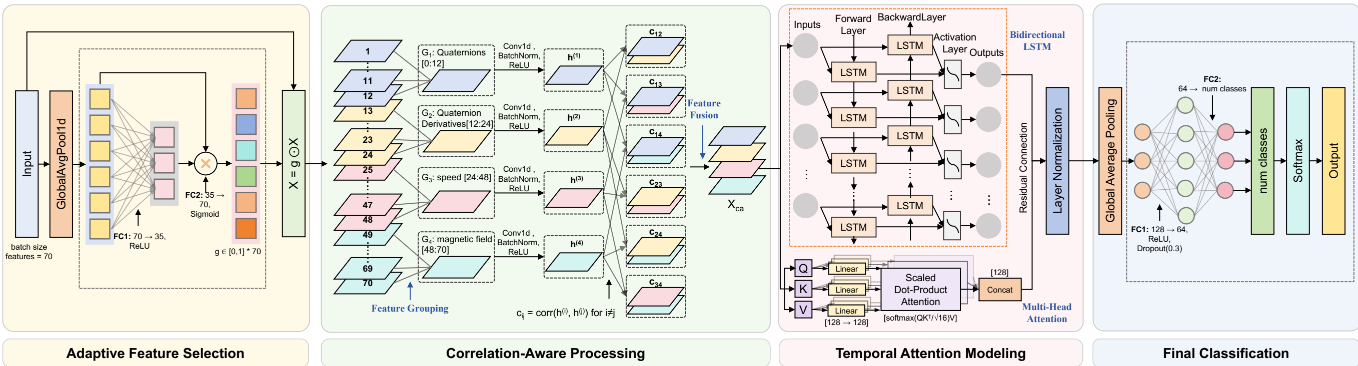

Fig. 1. **CareAttenNet Architecture**. An overview of the CareAttenNet architecture for nursing activity recognition. The framework processes sensor data through four sequential stages: (1) an Adaptive Feature Selection module to weight input channels, (2) a Correlation-Aware Processing stage to model inter-sensor relationships, (3) a Temporal Attention Modeling stage using a Bidirectional LSTM and Multi-Head Attention to capture temporal dependencies, (4) a Final Classification stage to output activity predictions.

types from start to end while obtaining verbal confirmation from nurses to ensure correct label selection, with subsequent validation enabled through synchronized pose estimation data from parallel video recordings.

C. Model Architectures

We propose CareAttenNet (Care Activity Recognition with Attention Network), a deep learning architecture integrating adaptive feature selection, correlation-aware processing, and temporal attention for nursing activity recognition as shown in Figure 1. We compare CareAttenNet against four baseline architectures to validate the effectiveness of our integrated approach.

CareAttenNet implements a three-stage processing pipeline addressing specific challenges in nursing activity recognition:

a) Stage 1: Adaptive Feature Selection: A two-layer gating network learns importance weights for input features:

$$\mathbf{g}_{fs} = \text{sigmoid}(\mathbf{W}_{fs} \tanh(\mathbf{W}'_{fs}\mathbf{X} + \mathbf{b}'_{fs}) + \mathbf{b}_{fs}) \tag{3}$$

Applied element-wise to suppress irrelevant channels:

$$\mathbf{X}_{fs} = \mathbf{g}_{fs} \odot \mathbf{X} \tag{4}$$

b) Stage 2: Correlation-Aware Processing: Computes batch-wise correlation matrices to model inter-sensor relationships:

$$\mathbf{R}_t = \text{BatchCorr}(\mathbf{X}_{fs}) \tag{5}$$

Processes original and correlation-weighted features jointly:

$$\mathbf{h}_{ca} = \text{Conv1D}([\mathbf{X}_{fs}, \mathbf{R}_t\mathbf{X}_{fs}]) \tag{6}$$

c) Stage 3: Temporal Attention Modeling: Bidirectional LSTM captures temporal context:

$$\mathbf{h}_{lstm} = \text{BiLSTM}(\mathbf{h}_{ca}) \tag{7}$$

Attention mechanism computes temporal importance weights:

$$e_t = \mathbf{v}^T \tanh(\mathbf{W}_e\mathbf{h}_{lstm,t} + \mathbf{b}_e) \tag{8}$$

$$\alpha_t = \frac{\exp(e_t)}{\sum_{j=1}^{T}\exp(e_j)} \tag{9}$$

Context vector aggregates temporal information:

$$\mathbf{c} = \sum_{t=1}^{T} \alpha_t \mathbf{h}_{lstm,t} \tag{10}$$

d) Multi-Scale Feature Fusion and Classification: Combines attention-weighted temporal representations with global features:

$$\mathbf{h}_{final} = \mathbf{W}_{fusion}[\mathbf{c}; \mathbf{h}_{lstm,T}; \text{GlobalAvgPool}(\mathbf{h}_{ca})] + \mathbf{b}_{fusion} \tag{11}$$

Final classification through softmax:

$$\mathbf{y} = \text{softmax}(\mathbf{W}_{clf}\mathbf{h}_{final} + \mathbf{b}_{clf}) \tag{12}$$

D. Baseline Architectures

We compare CareAttenNet against four baseline architectures:

CNN-LSTM: Standard hierarchical architecture with 1D convolutional layers (64, 128 filters, kernel size 3) and bidirectional LSTM (64 hidden units) [46].

Correlation-Aware CNN: Groups 70-dimensional inputs into four semantic categories (quaternion, derivatives, velocity, magnetic field) with inter-sensor correlation modeling [47].

Attention LSTM: Bidirectional LSTM with multi-head self-attention (8 heads) for temporal importance modeling [48].

Feature-Selective Network: Two-layer gating mechanism for adaptive sensor channel selection and noise reduction [49].

E. Training Methodology and Optimization

All models were trained using AdamW optimizer (learning rate: 1×10^{-4}, weight decay: 1×10^{-4}) with early stopping (patience: 100 epochs), ReduceLROnPlateau scheduling (factor: 0.5), and gradient clipping (max norm: 1.0). Cross-entropy loss with label smoothing ($\varepsilon = 0.1$) and batch size of 8 were employed (We experimented with batch sizes of 4, 8, and 16, and selected 8 based on optimal model performance and training stability.). To address class imbalance (156:1 ratio), we applied inverse frequency weighting, and used stratified data splitting. Models were evaluated using accuracy, weighted F1-score, precision, recall, ROC curves, and confusion matrices. This study employs non-overlapping temporal windows of 20 timesteps to segment sensor data, balancing sufficient temporal context for activity pattern recognition while preventing data leakage and ensuring each window contains a consistent activity label. The sensor features data were normalized before feeding into the model. Code is available at https://github.com/xiaoyanLi629/.

F. Ablation Study Design

We conducted comprehensive ablation studies to assess each architectural component's efficacy. Eight configurations were evaluated: CNN-LSTM [46], three individual components (feature selection mechanism, correlation awareness module, temporal attention), three pairwise combinations, and the complete CareAttenNet architecture. Identical training protocols and hyperparameters ensured fair comparison. Statistical testing validated performance improvements, quantifying individual contributions and synergistic interactions between components.

IV. Results

A. Data Statistics

The dataset contains 7,631,843 samples from 14 subjects performing 23 nursing activities. Each subject contributed 474,356 ± 321,927 samples across 16.9 ± 3.6 activities. Activity distribution (Figure 2D) analysis showed "change clothes" as most frequent (259,286 samples), followed by "clean up" (165,120) and "make bed" (137,428). All activities exceeded 1,000 samples, ensuring robust analysis. Activity durations varied significantly, from dental care (86,455 seconds average) to blow-drying (317,273 seconds). The 70-dimensional sensor features (Figure 2C) data included quaternion measurements, angular velocities, linear accelerations, and magnetic field readings. Correlation analysis revealed six feature pairs with $|r| > 0.8$ and 19 pairs with $|r| > 0.5$ (maximum 0.959), indicating potential redundancy and dimensionality reduction opportunities (Figure 2E). The UMAP plot in Figure 2F demonstrates that data points corresponding to the same activity are projected to similar spatial locations, resulting in the formation of well-defined clusters.

B. Experiment Results

The experimental results revealed varying levels of performance across different architectural configurations. CareAttenNet demonstrated superior performance, achieving 77.36% validation accuracy and 60.00% test accuracy, outperforming other architectures in both metrics as shown in Table I. The Attention LSTM [48] followed closely with 76.53% validation accuracy and 58.33% test accuracy, while the CNN-LSTM [46] achieved 75.96% validation accuracy and 57.92% test accuracy. The confusion matrix heatmap demonstrates CareAttenNet's classification performance across nursing activities, where "put accessories" and "serve food" show the strongest diagonal intensities indicating best recognition accuracy, while activities like "wipe up" and "wheelchair transfer" exhibit higher confusion rates, particularly between functionally similar tasks, suggesting challenges in distinguishing activities that share similar motion patterns or sensor signatures as shown in Figure 2G.

The Correlation-Aware CNN [47] showed relatively lower performance with 71.98% validation accuracy and 54.62% test accuracy, suggesting limited benefits from explicit sensor correlation modeling. The Feature-Selective Network [49] achieved 68.55% validation accuracy and 52.86% test accuracy, indicating that adaptive feature selection alone may not provide sufficient discriminative power for robust activity recognition.

Performance metrics across architectures showed consistent patterns between accuracy and F1-score, suggesting balanced performance despite class imbalance in the dataset. Precision metrics generally exceeded recall across all models, indicating conservative prediction behavior that prioritizes prediction confidence over coverage. CareAttenNet maintained the highest precision at 79.39% while achieving 77.36% recall on the validation set.

Ablation studies in Table II reveal temporal attention as the most effective single component (78.33% test accuracy). Feature selection and correlation-aware provide modest individual improvements (69.33% and 64.34%), while their combination yields mixed results. The complete CareAttenNet architecture shows unexpected degradation compared to simpler combinations, suggesting integration challenges. Feature + Attention achieves competitive performance (77.40%), indicating synergy between adaptive selection and temporal modeling, while Correlation + Attention shows strong validation but limited test generalization.

V. Discussion

In response to increasing healthcare system pressures, accurate identification and documentation of nursing activities are crucial for improving care quality and optimizing resource allocation.

A. Module Interaction Conflicts

Our experimental analysis reveals several unexpected challenges that provide important insights into the complexity of multi-modal sensor fusion for healthcare applications. The counterintuitive performance degradation when combining all CareAttenNet components (60.00% test accuracy) compared to Feature + Attention (77.40% test accuracy) reveals complex negative interactions that challenge modular architecture assumptions. The correlation-aware processing module computes batch-wise correlation matrices $\mathbf{R}t = \text{BatchCorr}(\mathbf{X}fs)$ while temporal attention learns importance weights through $\alpha_t = \frac{\exp(e_t)}{\sum_{j=1}^{T} \exp(e_j)}$, creating competing adaptive mechanisms that may interfere with each other. This suggests information

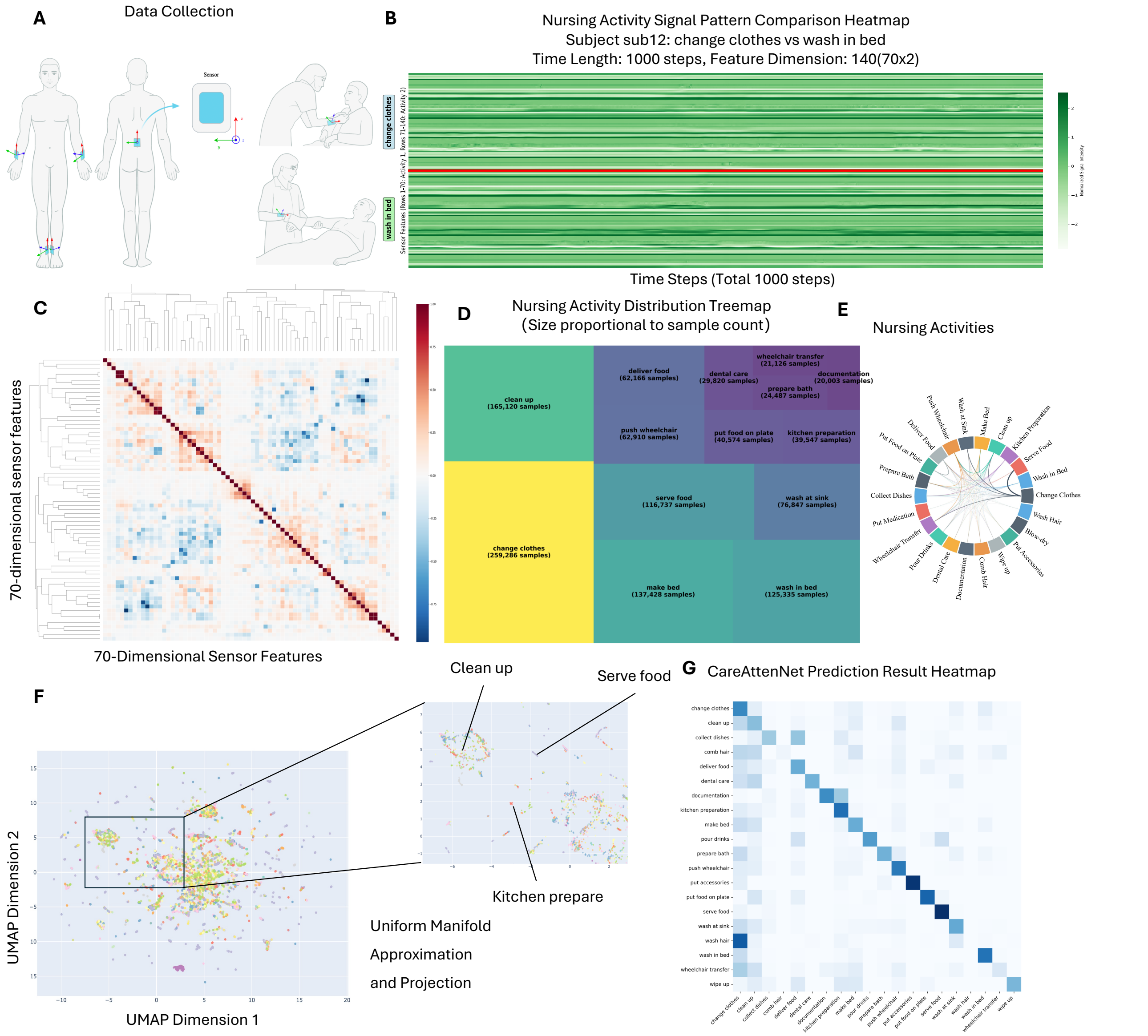

Fig. 2. **Dataset composition and statistical characterization**. **A**, Data acquisition protocol utilizing wearable inertial measurement units (IMUs) to capture sensor measurements from 14 healthcare professionals executing more than 20 distinct nursing activities in clinical environments. **B**, Representative temporal signal patterns for change clothes and wash in bed activities demonstrating characteristic motion signatures. **C**, Correlation matrix visualization of the 70-dimensional sensor feature space revealing inter-feature dependencies and redundancy patterns. **D**, Hierarchical treemap visualization depicting the relative frequency distribution of nursing activities within the dataset. **E**, Co-occurrence correlation analysis of 22 nursing activity categories selected for having over 1,000 samples each. **F**, UMAP dimensionality reduction projection illustrating the geometric structure and separability of high-dimensional activity representations. **G**, Confusion matrix heatmap quantifying CareAttenNet classification performance with predicted labels versus ground truth annotations.

redundancy between correlation processing and temporal attention mechanisms, where both capture sensor interdependencies through different pathways, resulting in optimization conflicts rather than enhanced performance.

B. Temporal Modeling Challenges

Beyond module interactions, the temporal modeling component faces fundamental challenges stemming from nursing activity variability. The extreme duration variations (86,455 to 317,273 seconds) expose limitations in CareAttenNet's fixed-window temporal modeling approach. The bidirectional LSTM architecture requires either truncation of long activities or padding of short activities, introducing systematic biases that compromise temporal pattern learning. Activities longer than window T reduce to local motion classification rather than comprehensive activity recognition, while shorter activities suffer from temporal dilution through zero-padding. This fixed-window approach fails to capture activity boundary dynamics and phase transitions crucial for distinguishing sequential activities, suggesting adaptive temporal windowing may be necessary.

TABLE I
PERFORMANCE COMPARISON OF CAREATTENNET AGAINST BASELINE ARCHITECTURES

Metrics	**Train**				**Validation**				**Test**			
	Acc	F1	Prec	Recall	Acc	F1	Prec	Recall	Acc	F1	Prec	Recall
CNN-LSTM	0.86	0.86	0.86	0.86	0.76	0.76	0.78	0.76	0.58	0.58	0.62	0.58
Correlation CNN	0.80	0.80	0.80	0.80	0.72	0.72	0.75	0.72	0.55	0.55	0.58	0.55
Attention LSTM	**0.89**	**0.89**	0.89	**0.89**	**0.77**	0.77	**0.79**	**0.77**	0.58	**0.59**	0.62	0.58
Feature SelectiveNet	0.76	0.76	0.77	0.76	0.69	0.69	0.71	0.69	0.53	0.53	0.55	0.53
CareAttenNet	0.88	0.88	**0.89**	0.88	**0.77**	**0.78**	**0.79**	**0.77**	**0.60**	**0.60**	**0.64**	**0.60**

TABLE II
ABLATION STUDY RESULTS SHOWING INDIVIDUAL AND COMBINED CONTRIBUTIONS OF CAREATTENNET COMPONENTS

Model	**Train**				**Validation**				**Test**			
	Acc	F1	Prec	Recall	Acc	F1	Prec	Recall	Acc	F1	Prec	Recall
Baseline	0.78	0.78	0.79	0.78	0.33	0.34	0.40	0.33	0.68	0.68	0.70	0.68
Feature Selection Only	0.80	0.80	0.80	0.80	0.35	0.35	0.41	0.35	0.69	0.69	0.71	0.69
Correlation Aware Only	0.73	0.73	0.74	0.74	0.35	0.34	0.37	0.35	0.64	0.64	0.67	0.64
Temporal Attention Only	**0.90**	**0.90**	**0.90**	**0.90**	0.39	0.40	0.45	0.39	**0.78**	**0.79**	**0.81**	0.78
Feature + Correlation	0.78	0.77	0.78	0.78	0.33	0.33	0.39	0.33	0.67	0.67	0.68	0.67
Feature + Attention	0.89	0.89	0.89	0.89	0.37	0.37	0.43	0.37	0.77	0.78	0.80	0.77
Correlation + Attention	0.89	0.89	0.89	0.89	**0.77**	**0.78**	**0.79**	**0.77**	0.55	0.54	0.60	0.55
Full CareAttenNet	0.88	0.88	0.89	0.88	**0.77**	**0.78**	**0.79**	**0.77**	0.60	0.60	0.64	0.60

C. Physics-Informed Feature Grouping Limitations

Complementing these architectural challenges, our investigation into physics-informed feature representations reveals additional complexity in domain knowledge integration. The suboptimal performance of physics-informed grouping (54.62% test accuracy) suggests that our quaternion-velocity-magnetic field decomposition may be insufficient to capture the complete spectrum of biomechanical relationships underlying nursing activities. The complex motor patterns in healthcare tasks likely involve higher-order physical interactions—such as multi-joint coordination dynamics, force distribution patterns, or vestibular-proprioceptive coupling—that are not adequately represented by the current decomposition. This limitation suggests that more comprehensive physics-based modeling is needed, potentially incorporating additional sensor modalities or more sophisticated relationship modeling that captures multi-scale biomechanical dependencies without over-constraining the learning process.

VI. CONCLUSION

This study presents CareAttenNet model, a novel deep learning architecture for automated nursing activity recognition that achieved 60.00% test accuracy on the SONAR dataset, outperforming baseline methods. Temporal attention emerged as the most effective component (78.33% individual accuracy), while combining all architectural components unexpectedly degraded performance, revealing complex negative interactions between correlation processing and temporal attention mechanisms.

The analysis exposed fundamental challenges in healthcare sensor fusion, including limitations of fixed-window approaches for variable-duration activities (86,455 to 317,273 seconds) and insufficient physics-informed feature grouping. These findings highlight the need for adaptive temporal windowing, enhanced biomechanical modeling, and systematic investigation of component integration challenges.

Future research should focus on cross-domain generalization, real-time processing capabilities, and clinical deployment considerations. This work establishes a foundation for intelligent healthcare monitoring systems that could reduce documentation burden and support nursing professionals, while demonstrating the nuanced challenges of applying AI to complex healthcare environments. While our current approach computes batch-wise correlation matrices, exploring running correlations or temporal correlation measures (such as sliding window correlations or time-lagged cross-correlations) that could better capture the dynamic relationships in nursing activities.

ACKNOWLEDGMENT

This work was supported by the Young Scientists Fund of the Natural Science Foundation of Shandong Province (ZR2022QF134).

REFERENCES

[1] Linda H Aiken, Douglas M Sloane, Luk Bruyneel, Koen Van den Heede, Peter Griffiths, Reinhard Busse, Marianna Diomidous, Juha Kinnunen, Maria Kózka, Emmanuel Lesaffre, et al. Nurse staffing and education and

hospital mortality in nine european countries: a retrospective observational study. *The lancet*, 383(9931):1824–1830, 2014.

[2] Katarina Edfeldt, Lena Nyholm, Eva Jangland, Anna-Karin Gunnarsson, Camilla Fröjd, and Anna Hauffman. Missed nursing care in surgical care–a hazard to patient safety: a quantitative study within the incharge programme. *BMC nursing*, 23(1):233, 2024.

[3] Hyoil Han, Yoori Choi, Yoo Myung Choi, Xiaohua Zhou, and Ari D Brooks. A generic framework: From clinical notes to electronic medical records. In *19th IEEE Symposium on Computer-Based Medical Systems (CBMS'06)*, pages 111–118. IEEE, 2006.

[4] Elske Ammenwerth and H-P Spötl. The time needed for clinical documentation versus direct patient care. *Methods of information in medicine*, 48(01):84–91, 2009.

[5] Anjanava Biswas and Wrick Talukdar. Intelligent clinical documentation: Harnessing generative ai for patient-centric clinical note generation. *arXiv preprint arXiv:2405.18346*, 2024.

[6] Parisa Rashidi and Alex Mihailidis. A survey on ambient-assisted living tools for older adults. *IEEE journal of biomedical and health informatics*, 17(3):579–590, 2012.

[7] Urvashi Tandon, Myriam Ertz, and Shashi. Continued intention of mhealth care applications among the elderly: an enabler and inhibitor perspective. *International Journal of Human–Computer Interaction*, 40(18):5026–5041, 2024.

[8] Attila Reiss and Didier Stricker. Introducing a new benchmarked dataset for activity monitoring. In *2012 16th international symposium on wearable computers*, pages 108–109. IEEE, 2012.

[9] Oscar D Lara and Miguel A Labrador. A survey on human activity recognition using wearable sensors. *IEEE communications surveys & tutorials*, 15(3):1192–1209, 2012.

[10] Harish Haresamudram, Chi Ian Tang, Sungho Suh, Paul Lukowicz, and Thomas Ploetz. Past, present, and future of sensor-based human activity recognition using wearables: A surveying tutorial on a still challenging task. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 9(2):1–44, 2025.

[11] Yuqing Chen and Yang Xue. A deep learning approach to human activity recognition based on single accelerometer. In *2015 IEEE international conference on systems, man, and cybernetics*, pages 1488–1492. IEEE, 2015.

[12] Charissa Ann Ronao and Sung-Bae Cho. Human activity recognition with smartphone sensors using deep learning neural networks. *Expert systems with applications*, 59:235–244, 2016.

[13] Jing Liu, Wei Zhu, Di Li, Xing Hu, and Liang Song. Domain generalization with semi-supervised learning for people-centric activity recognition. *Science China Information Sciences*, 68(1):112103, 2025.

[14] Enrique Garcia-Ceja, Michael Riegler, Tine Nordgreen, Petter Jakobsen, Ketil J Oedegaard, and Jim Tørresen. Mental health monitoring with multimodal sensing and machine learning: A survey. *Pervasive and Mobile Computing*, 51:1–26, 2018.

[15] Parisa Rashidi, Diane J Cook, Lawrence B Holder, and Maureen Schmitter-Edgecombe. Discovering activities to recognize and track in a smart environment. *IEEE transactions on knowledge and data engineering*, 23(4):527–539, 2010.

[16] Sebastian OH Madgwick, Andrew JL Harrison, and Ravi Vaidyanathan. Estimation of imu and marg orientation using a gradient descent algorithm. In *2011 IEEE international conference on rehabilitation robotics*, pages 1–7. Ieee, 2011.

[17] Jinliang Zang, Le Wang, Ziyi Liu, Qilin Zhang, Gang Hua, and Nanning Zheng. Attention-based temporal weighted convolutional neural network for action recognition. In *Artificial Intelligence Applications and Innovations: 14th IFIP WG 12.5 International Conference, AIAI 2018, Rhodes, Greece, May 25–27, 2018, Proceedings 14*, pages 97–108. Springer, 2018.

[18] Christian Szegedy, Vincent Vanhoucke, Sergey Ioffe, Jon Shlens, and Zbigniew Wojna. Rethinking the inception architecture for computer vision. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 2818–2826, 2016.

[19] Richard Meyes, Melanie Lu, Constantin Waubert de Puiseau, and Tobias Meisen. Ablation studies in artificial neural networks. *arXiv preprint arXiv:1901.08644*, 2019.

[20] Orhan Konak, Valentin Döring, Tobias Fiedler, Lucas Liebe, Leander Masopust, Kirill Postnov, Franz Sauerwald, Felix Treykorn, Alexander Wischmann, Stefan Kalabakov, et al. Sonar, a nursing activity dataset with inertial sensors. *Scientific data*, 10(1):727, 2023.

[21] Sidney Katz. Assessing self-maintenance: activities of daily living, mobility, and instrumental activities of daily living. *Journal of the American Geriatrics Society*, 31(12):721–727, 1983.

[22] Bertrand Guidet, Helene Vallet, Hans Flaatten, Gavin Joynt, Sean M Bagshaw, Susannah K Leaver, Michael Beil, Bin Du, Daniel N Forte, Derek C Angus, et al. The trajectory of very old critically ill patients. *Intensive Care Medicine*, 50(2):181–194, 2024.

[23] Florence I Mahoney and Dorothea W Barthel. Functional evaluation: the barthel index: a simple index of independence useful in scoring improvement in the rehabilitation of the chronically ill. *Maryland state medical journal*, 1965.

[24] Shylie Mackintosh. *The Functional Independence Measure*. PhD thesis, Australian Physiotherapy Association, 2009.

[25] Julie Bernhardt, Kathryn S Hayward, Gert Kwakkel, Nick S Ward, Steven L Wolf, Karen Borschmann, John W Krakauer, Lara A Boyd, S Thomas Carmichael, Dale Corbett, et al. Agreed definitions and a shared vision for new standards in stroke recovery research: the

stroke recovery and rehabilitation roundtable taskforce. *International Journal of Stroke*, 12(5):444–450, 2017.

[26] Mary E Tinetti. Performance-oriented assessment of mobility problems in elderly patients. *Journal of the American Geriatrics Society*, 1986.

[27] Lisa Sheehy, Lalita Bharadwaj, Kelsey Annie Nissen, and Justine L Estey. Non-immersive virtual reality exercise can increase exercise in older adults living in the community and in long-term care: A randomized controlled trial. *Clinical Interventions in Aging*, pages 109–124, 2025.

[28] Diane Podsiadlo and Sandra Richardson. The timed "up & go": a test of basic functional mobility for frail elderly persons. *Journal of the American geriatrics Society*, 39(2):142–148, 1991.

[29] JM Mores. Preventing patient falls. establishing a fall intervention program, 2009.

[30] Jialuo Du, Yidong Ren, Zhui Zhu, Chenning Li, Zhichao Cao, Qiang Ma, and Yunhao Liu. Srlora: Neural-enhanced lora weak signal decoding with multi-gateway super resolution. In *Proceedings of the Twenty-fourth International Symposium on Theory, Algorithmic Foundations, and Protocol Design for Mobile Networks and Mobile Computing*, pages 270–279, 2023.

[31] Fei Jiang, Yong Jiang, Hui Zhi, Yi Dong, Hao Li, Sufeng Ma, Yilong Wang, Qiang Dong, Haipeng Shen, and Yongjun Wang. Artificial intelligence in healthcare: past, present and future. *Stroke and vascular neurology*, 2(4), 2017.

[32] Riccardo Miotto, Fei Wang, Shuang Wang, Xiaoqian Jiang, and Joel T Dudley. Deep learning for healthcare: review, opportunities and challenges. *Briefings in bioinformatics*, 19(6):1236–1246, 2018.

[33] Kun-Hsing Yu, Andrew L Beam, and Isaac S Kohane. Artificial intelligence in healthcare. *Nature biomedical engineering*, 2(10):719–731, 2018.

[34] Andre Esteva, Alexandre Robicquet, Bharath Ramsundar, Volodymyr Kuleshov, Mark DePristo, Katherine Chou, Claire Cui, Greg Corrado, Sebastian Thrun, and Jeff Dean. A guide to deep learning in healthcare. *Nature medicine*, 25(1):24–29, 2019.

[35] Jianxing He, Sally L Baxter, Jie Xu, Jiming Xu, Xingtao Zhou, and Kang Zhang. The practical implementation of artificial intelligence technologies in medicine. *Nature medicine*, 25(1):30–36, 2019.

[36] Andrew L Beam and Isaac S Kohane. Big data and machine learning in health care. *Jama*, 319(13):1317–1318, 2018.

[37] Alvin Rajkomar, Jeffrey Dean, and Isaac Kohane. Machine learning in medicine. *New England Journal of Medicine*, 380(14):1347–1358, 2019.

[38] Sebastian Vollmer, Bilal A Mateen, Gergo Bohner, Franz J Király, Rayid Ghani, Pall Jonsson, Sarah Cumbers, Adrian Jonas, Katherine SL McAllister, Puja Myles, et al. Machine learning and artificial intelligence research for patient benefit: 20 critical questions on transparency, replicability, ethics, and effectiveness. *bmj*, 368, 2020.

[39] Travers Ching, Daniel S Himmelstein, Brett K Beaulieu-Jones, Alexandr A Kalinin, Brian T Do, Gregory P Way, Enrico Ferrero, Paul-Michael Agapow, Michael Zietz, Michael M Hoffman, et al. Opportunities and obstacles for deep learning in biology and medicine. *Journal of the royal society interface*, 15(141):20170387, 2018.

[40] Min Chen, Shiwen Mao, and Yunhao Liu. Big data: A survey. *Mobile networks and applications*, 19(2):171–209, 2014.

[41] Kee Yuan Ngiam and Wei Khor. Big data and machine learning algorithms for health-care delivery. *The Lancet Oncology*, 20(5):e262–e273, 2019.

[42] Pratik Shah, Francis Kendall, Sean Khozin, Ryan Goosen, Jianying Hu, Jason Laramie, Michael Ringel, and Nicholas Schork. Artificial intelligence and machine learning in clinical development: a translational perspective. *NPJ digital medicine*, 2(1):69, 2019.

[43] Oliver Faust, Yuki Hagiwara, Tan Jen Hong, Oh Shu Lih, and U Rajendra Acharya. Deep learning for healthcare applications based on physiological signals: A review. *Computer methods and programs in biomedicine*, 161:1–13, 2018.

[44] E Sandeep Kumar and Pappu Satya Jayadev. Deep learning for clinical decision support systems: a review from the panorama of smart healthcare. *Deep learning techniques for biomedical and health informatics*, pages 79–99, 2019.

[45] Benjamin Shickel, Patrick James Tighe, Azra Bihorac, and Parisa Rashidi. Deep ehr: a survey of recent advances in deep learning techniques for electronic health record (ehr) analysis. *IEEE journal of biomedical and health informatics*, 22(5):1589–1604, 2017.

[46] Yuan Tao, Chao Liu, Tianyang Chen, Xingwang Zhao, Chunyang Liu, Haojie Hu, Tengfei Zhou, and Haiqiang Xin. Real-time multipath mitigation in multi-gnss short baseline positioning via cnn-lstm method. *Mathematical Problems in Engineering*, 2021(1):6573230, 2021.

[47] Arvin Hekmati, Jiahe Zhang, Tamoghna Sarkar, Nishant Jethwa, Eugenio Grippo, and Bhaskar Krishnamachari. Correlation-aware neural networks for ddos attack detection in iot systems. *IEEE/ACM Transactions on Networking*, 32(5):3929–3944, 2024.

[48] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, and Illia Polosukhin. Attention is all you need. *Advances in neural information processing systems*, 30, 2017.

[49] Haojie Ma, Wenzhong Li, Xiao Zhang, Songcheng Gao, and Sanglu Lu. Attnsense: Multi-level attention mechanism for multimodal human activity recognition. In *IJCAI*, pages 3109–3115, 2019.