engineering proceedings

Proceeding Paper

Machine Learning-Based Posture Recognition Architecture for an Intelligent Medical Restraint Glove †

Yen-Chieh Mao [1,*] and Chia-An Yen [2]

[1] Department of Mechanical Design Engineering, National Formosa University, Yunlin 63201, Taiwan
[2] Department of Precision Mechanical Engineering, National Formosa University, Yunlin 63201, Taiwan; yenjia1104@gmail.com
* Correspondence: yjmau63@gmail.com
† Presented at the 6th International Conference on Electronic Communications, Internet of Things and Big Data (ICEIB 2026), New Taipei, Taiwan, 24–26 April 2026.

Abstract

We developed an IoT–enabled edge intelligence system to detect unplanned nasogastric tube removal in clinical care using a smart restraint glove. Real-time hand gesture recognition is achieved through a tiny machine learning (TinyML) deep learning architecture that integrates convolutional neural networks (CNN) and long short-term memory networks. Deployed on an Espressif Systems 32 microcontroller, the system enables low-latency on-device inference. Transfer learning addresses limited labeled data by pre-training CNN layers on public human activity datasets, then fine-tuning with medical gesture signals from an MPU6050 inertial measurement unit. Local inference enhances privacy while IoT alerts notify staff.

Keywords: TinyML; transfer learning; deep learning; edge computing

1. Introduction

With the aging of global populations and increasing reliance on invasive medical devices in acute care, a number of patients rely on nasogastric tubes, urinary catheters and/or endotracheal tubes to sustain physiological function and therapeutic efficacy. In clinical practice, however, altered consciousness, discomfort, and agitation render patients susceptible to unplanned extubation (UE). The median UE incidence among adult inpatients is approximately 7.3%, with variation across study designs and clinical environments [1]. Most UE-affected patients require emergency re-intubation, a procedure usually associated with airway injury, cardiovascular instability, prolonged hospitalization, and elevated mortality risk [1,2]. Each UE event additionally imposes considerable economic costs and is linked to an increase in in-hospital mortality [1]. Their secondary complications, such as aspiration pneumonia, compound the nursing workload and systemic healthcare burden. Real-time detection and early warning of patient self-extubation behavior represents a clinically significant problem in intelligent healthcare technology developments.

Current preventive strategies against UE fall into two categories: mechanical restraint and electronic monitoring. Mechanical restraints, including wrist straps, restraint gloves, and bed rails, are the most widely adopted interventions. Souza et al. demonstrated that sustained physical restriction raises ethical concerns and is associated with patient anxiety, agitation, pressure injuries, and circulatory impairment. Camera-based electronic monitoring systems offer an alternative [3]. Gabriel et al. highlighted the potential of AI to augment such surveillance; however, these systems are susceptible to occlusion from

Check for updates

Academic Editors: Teen-Hang Meen, Shu-Han Liao and Cheng-Fu Yang

Published: 19 November 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

patient repositioning and bedding, and continuous video recording raises serious patient privacy concerns [4].

To address the comfort limitations of mechanical restraint and the privacy risks of visual monitoring, we developed a machine learning-based hand gesture recognition system embedded within a wearable medical restraint glove. The system is built on an Espressif Systems microcontroller (ESP32 (Espressif Systems, Shanghai, China)) integrated with an MPU6050 (TDK InvenSense, Kanagawa, Japan) six-axis inertial measurement unit (IMU) and tiny machine learning (TinyML) inference engine, acquiring real-time multi-axis acceleration and angular velocity signals to capture wrist rotation and fine-grained motion characteristics. The sensing module is compact and low power, and the system performs on-device inference without persistent cloud connectivity, enabling real-time UE behavior detection while minimizing patient privacy exposure.

2. Related Works

2.1. Wearable Human Activity Recognition (HAR)

HAR technology is widely applied in smart healthcare, rehabilitation monitoring, and wearable systems, using IMU-based sensors such as accelerometers and gyroscopes to capture motion signals. Early studies relied on conventional classifiers, such as support vector machines, decision trees, and random forests, paired with manual feature engineering over statistical or frequency-domain representations, proving difficulties in generalizing to complex temporal motion patterns [5,6]. Deep learning techniques shifted this paradigm. For example, convolutional neural network (CNN) and recurrent neural network (RNN) architectures allow models to extract high-level features directly from raw sensor data, and hybrid CNN-long short-term memory (LSTM) models in particular demonstrate a high recognition performance across multiple benchmark datasets by jointly capturing spatial and temporal characteristics [7–10].

2.2. Medical Data Scarcity and Transfer Learning in HAR

In medicine, the high cost of data acquisition and annotation leads to a situation of chronic disease sample insufficiency, raising overfitting risk and limiting model generalizability [11]. Transfer learning addresses this by pre-training on a different data-rich source domain to acquire generalizable representations, then fine-tuning on the target domain under constrained data conditions [12–14]. Complementary strategies such as few-shot learning and multi-source training have further improved accuracy and robustness [15]. In this study, we used the OPPORTUNITY Activity Recognition Dataset as the pre-training source [16] owing to its diverse coverage of daily activities and hand manipulation behaviors [17]. Its kinematic profiles share meaningful commonality with the clinical target taskand have validated training and recognition performance through prior deep learning studies [10,18].

2.3. TinyML and Edge Computing

Recent advances in IoT and embedded systems have made it practical to realize an on-device inference system. TinyML enables lightweight deep learning models to execute on microcontrollers without consistent cloud dependency, supporting real-time inference under tight power and memory systems [12,19]. CNN and LSTM architectures have been deployed on edge and wearable platforms through model quantization and architectural optimization, achieving low-latency recognition with reduced computational overhead [20–22]. For medical applications specifically, edge deployment also reduces data transmission exposure, which is a critical privacy consideration.

https://doi.org/10.3390/engproc2026159050

2.4. Limitations of Current Medical Restraint Devices

Existing UE prevention strategies rely mainly on physical restraints such as gloves and fixation devices, which lack real-time behavioral monitoring or intent prediction. Recent studies have explored sensor-based and AI–driven approaches to improve patient behavior detection [23,24]. However, a practical embedded solution that integrates inertial IMU sensing, deep learning inference, and edge deployment within a clinical restraint device remains absent. This study addresses that gap by proposing a unified system that combines IMU data acquisition, advanced neural architectures, and edge computing to enable proactive monitoring and intervention in clinical environments.

3. Methodology

3.1. Architecture

The developed adopts a honeycomb-structured restraint glove as the wearable carrier, as shown in Figure 1. The three-dimensional hexagonal lattice geometry reduces overall glove weight and promotes ventilation, improving long-duration wearing comfort and reducing skin loading, while providing a stable internal cavity to accommodate the ESP32, IMU, and associated electronic components.

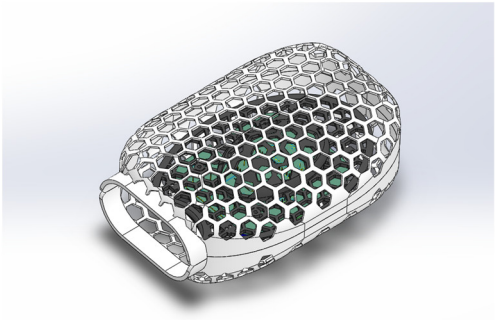


Figure 1. The design of the proposed constraint glove.

The electric control module within the glove body includes said IMU, ESP32, and a battery and charging/discharging control unit, enabling real-time data acquisition, on-device processing, and edge inference (Figure 2). This design is lightweight, consumes low power, and confirms long-term wearing comfort, fitting the basic clinical deployment requirements.

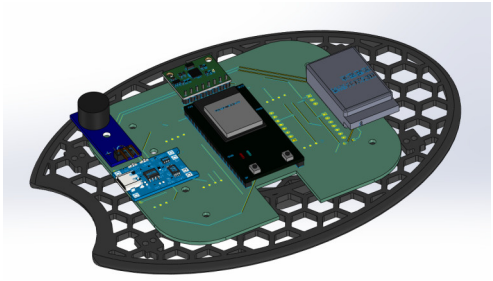


Figure 2. The electric control module of the proposed restraint glove.

The ESP32 microcontroller serves as the computing unit, responsible for sensor data acquisition, pre-processing, and deep learning model inference. The MPU6050 sensor, located beneath the center of the wearer's palm, is adopted as the sensing module to measure three-axis accelerations and three-axis angular velocities of the hand motion.

https://doi.org/10.3390/engproc2026159050

3.2. Source Dataset Selection and Preprocessing

The quality and quantity of the source dataset affect the feature extraction performance of the pre-trained model for transfer learning. We used the OPPORTUNITY Activity Recognition Dataset as the source domain [16], a well-established benchmark in the HAR field that encompasses diverse daily activities and fine-grained hand manipulation behaviors across multi-sensor, multi-dimensional recordings [17].

To align with the tube-removal recognition objective of this study, the original multi-class labels are restructured into a binary classification problem. General daily activities and low-risk movements are categorized as Non-Danger, while behaviors involving facial approach or manual pulling actions are categorized as Danger. Time-series data segmentation employs a sliding window mechanism with inter-epoch gap checking to improve sensitivity to high-risk motion patterns.

3.3. Data Preprocessing

The target dataset was collected using this restraint glove in a simulated clinical environment, comprising hand motion samples representative of pre-extubation behaviors, including hand-raising, grasping, and tube-pulling actions. Sensor data were sampled at 10 Hz; each sample consists of three-axis accelerations and three-axis angular velocities. Three motion classes are defined: static stable state, activities of daily living, and tube-removal risk state, corresponding respectively to resting, routine movement, and extubation-risk behaviors. Raw time-series data are segmented in a sliding window of length $T = 50$ frames and stride $s = 10$ frames. The i-th window slice is defined as:

$$\mathbf{X}^{(i)} = [x_{i-s}, x_{i-s+1}, \dots, x_{i-s+T-1}] \in \mathbf{R}^{\mathbf{T}\times 6} \tag{1}$$

Here, $\mathbf{X}^{(i)}$ denotes the i-th window, $x_t \in \mathbf{R}^6$ representing the six-axis input at time t, and the label is determined by the last frame to reflect the end-of-window action state.

3.4. Deep Time-Series Model Architecture

The developed system combines CNN and LSTM to extract spatial and temporal features from IMU sensor data. The overall architecture comprises a recency weighting layer, a CNN spatial feature extraction module, an LSTM temporal modeling layer, and a classification output layer, designed to enhance recognition of the tube-removal risk state and to support cross-domain transfer learning.

3.4.1. Recency Weight

To emphasize the action-critical information in the latter portion of each window, a Recency Weight Layer is introduced at the input stage, applying linearly increasing weights across time steps:

$$w_t = w_{\min} + \frac{(w_{\max} - w_{\min}) \cdot t}{T - 1} \tag{2}$$

Here, $w_{\max}$ and $w_{\min}$ denote the maximum and minimum weights, respectively.

3.4.2. LSTM and Classification

The LSTM module operation is summarized as:

$$h_t, c_t = \mathrm{LSTM}(x_t, h_{t-1}, c_{t-1}) \tag{3}$$

where x_t represents the input feature sequence derived from CNN and pooling operations, h_{t-1} and c_{t-1} denote the hidden state and cell state from the previous time step, h_t and c_t are the corresponding updated states at the current time step.

https://doi.org/10.3390/engproc2026159050

4. Results and Discussion

4.1. Pre-Training Performance on Source Dataset

During the pre-training phase on the OPPORTUNITY dataset, binary classification between danger and non-danger motion classes was evaluated, with results summarized in Table 1. The "Support" column indicates the actual number of samples for each category in the test set. The model demonstrates effective recognition of both low-risk general movements and high-risk raising actions, achieving an overall classification accuracy of 95.00%. The danger class recall reaches 0.9812, indicating that the pre-trained model is highly sensitive and capable of detecting nearly all danger samples. Although the danger class F1-score is 0.8698, the elevated recall confirms that the model has successfully learned discriminative high-risk temporal features, establishing a robust pre-training foundation for subsequent transfer learning under more complex conditions.

Table 1. Performance of the pre-trained binary classification model.

Category	Precision	Recall	F1-Score	Support
Non-danger	0.9959	0.9436	0.9691	780
Danger	0.7811	0.9812	0.8698	160
Accuracy			0.9500	940
Average	0.8885	0.9624	0.9194	940
Weighted average	0.9594	0.9500	0.9522	940

4.2. Transfer Learning Performance on Target Dataset

The three-class classification results on the IMU target dataset are summarized in Table 2, with the overall accuracy reaching 97.78%. Recognition of the Tube-Removal Risk State is particularly strong, with a recall and F1-score of 0.9872 and 0.9923, respectively, demonstrating that the model accurately captures clinically meaningful high-risk motion characteristics.

Table 2. Performance on IMU after fine-tuning for datasets.

Category	Precision	Recall	F1-Score	Support
Static Stable State	0.9656	0.9873	0.9763	1024
Activities of Daily Living	0.9835	0.9634	0.9734	930
Tube-Removal Risk State	0.9974	0.9872	0.9923	390
Accuracy			0.9778	2344
Macro Avg	0.9822	0.9793	0.9807	2344
Weighted Avg	0.9780	0.9778	0.9778	2344

Examining per-class results, the Static Stable State achieves the highest recognition consistency, attributable to its low intra-class motion variability and stable feature distribution. The activities of daily living class presents greater recognition difficulty due to the inherent diversity of daily hand operations, which increases intra-class variance. Notably, the tube-removal risk state maintains high precision and recall simultaneously, confirming the model's ability to reliably distinguish high-risk actions from routine daily behavior and demonstrating strong potential for clinical deployment.

5. Conclusions

We developed a machine learning-based gesture recognition framework for a medical restraint glove, integrating a wearable sensing system with an embedded edge-computing platform, achieving real-time monitoring and prediction of self-extubation behaviors. The system performs data acquisition, time-series motion recognition, and alert generation

independently on-device without reliance on external computing infrastructure, reducing interference with patient daily activity and improving both clinical feasibility and patient privacy. The CNN-LSTM temporal model, pre-trained on a public dataset and adapted via transfer learning, achieves 94.47% accuracy on the binary source-domain classification task with a Danger class recall of 0.9938. On the three-class target task using IMU sensor data, the system attains an overall accuracy of 97.78%, with the tube-removal risk state reaching a recall of 0.9872 and an F1-score of 0.9923, confirming strong recognition capability for high-risk motion events. The developed transfer learning system, validated through edge deployment, demonstrates practical feasibility for wearable medical monitoring and real-time UE detection. Future work will focus on model compression and inference optimization, including quantization-aware training and data augmentation techniques, for patient safety monitoring and advancing edge AI applications in smart healthcare fields.

Author Contributions: Conceptualization, Y.-C.M.; methodology, Y.-C.M. and C.-A.Y.; software, C.-A.Y.; validation, C.-A.Y.; formal analysis, Y.-C.M. and C.-A.Y.; investigation, Y.-C.M. and C.-A.Y.; resources, Y.-C.M.; data curation, Y.-C.M. and C.-A.Y.; writing—original draft preparation, C.-A.Y.; writing—review and editing, Y.-C.M.; visualization, Y.-C.M. and C.-A.Y.; supervision, Y.-C.M.; project administration, Y.-C.M.; funding acquisition, Y.-C.M. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available upon request from the corresponding author.

Acknowledgments: The authors would like to acknowledge Cheng-Chi Ting, Fu-Hui Zhou, and Chia-Chuan Ho for their assistance with prototyping and data collection. During the preparation of this manuscript/study, the author(s) used Claude Sonnet 4.6 (Anthropic, 2026) for part of this article's English translation. The authors have reviewed and edited the output and take full responsibility for the content of this publication.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Berkow, L.; Kanowitz, A. Unplanned Extubation in the Perioperative Environment. 2022. Available online: https://www.apsf.org/article/unplanned-extubation-in-the-perioperative-environment/ (accessed on 14 April 2026).
2. de Lassence, A.; Alberti, C.; Azoulay, É.; Le Miere, E.; Cheval, C.; Vincent, F.; Cohen, Y.; Garrouste-Orgeas, M.; Adrie, C.; Troche, G.; et al. Impact of unplanned extubation and reintubation after weaning on nosocomial pneumonia risk in the intensive care unit: A prospective multicenter study. *Anesthesiology* **2002**, *97*, 148–156. [CrossRef] [PubMed]
3. da Silva Souza, L.M.; Santana, R.F.; da Silva Gabriel Capeletto, C.; Menezes, A.K.; Delvalle, R. Factors associated with mechanical restraint in the hospital environment: A cross-sectional study. *Rev. Esc. Enferm. USP* **2019**, *53*. [CrossRef] [PubMed]
4. Gabriel, P.; Rehani, P.; Troy, T.; Wyatt, T.; Choma, M.; Singh, N. Continuous patient monitoring with AI: Real-time analysis of video in hospital care settings. *Front. Imaging* **2025**, *4*, 1547166. [CrossRef]
5. Demrozi, F.; Pravadelli, G.; Bihorac, A.; Rashidi, P. Human activity recognition using inertial, physiological and environmental sensors: A comprehensive survey. *IEEE Access* **2020**, *8*, 210816–210836. [CrossRef] [PubMed]
6. Kaseris, M.; Kostavelis, I.; Malassiotis, S. A comprehensive survey on deep learning methods in human activity recognition. *Mach. Learn. Knowl. Extr.* **2024**, *6*, 842–876. [CrossRef]
7. Wang, Y.; Cang, S.; Yu, H. A survey on wearable sensor modality centred human activity recognition in health care. *Expert Syst. Appl.* **2019**, *137*, 167–190. [CrossRef]
8. Kolkar, R.; Singh Tomar, R.P.; Vasantha, G. IoT-based human activity recognition models based on CNN, LSTM and GRU. In Proceedings of the IEEE Silchar Subsection Conference (SILCON), Silchar, India, 4–6 November 2022; pp. 1–7. [CrossRef]
9. Hammerla, N.Y.; Halloran, S.; Plötz, T. Deep, convolutional, and recurrent models for human activity recognition using wearables. *arXiv* **2016**, arXiv:1604.08880. [CrossRef]

10. Ordóñez, F.J.; Roggen, D. Deep convolutional and LSTM recurrent neural networks for multimodal wearable activity recognition. *Sensors* **2016**, *16*, 115. [CrossRef] [PubMed]
11. Power, A.; Burda, Y.; Edwards, H.; Babuschkin, I.; Misra, V. Grokking: Generalization beyond overfitting on small algorithmic datasets. *arXiv* **2022**, arXiv:2201.02177. [CrossRef]
12. Hayajneh, A.M.; Hafeez, M.; Zaidi, S.A.R.; McLernon, D. TinyML empowered transfer learning on the edge. *IEEE Open J. Commun. Soc.* **2024**, *5*, 1656–1672. [CrossRef]
13. Ashry, S.; Das, S.; Rafiei, M.; Baumbach, J.; Baumbach, L. Transfer learning of human activities based on IMU sensors: A review. *IEEE Sens. J.* **2025**, *25*, 4115–4126. [CrossRef]
14. Abdulazeem, Y.; Balaha, H.M.; Bahgat, W.M.; Badawy, M. Human action recognition based on transfer learning approach. *IEEE Access* **2021**, *9*, 82058–82069. [CrossRef]
15. Fan, T.; Qiu, S.; Gong, W.; Fang, Y. Multi-source domain generalization for CSI-based human activity recognition. *IEEE Trans. Mob. Comput.* **2025**, *24*, 11034–11045. [CrossRef]
16. Roggen, D.; Calatroni, A.; Nguyen-Dinh, L.V.; Chavarriaga, R.; Sagha, H. *OPPORTUNITY Activity Recognition Dataset*; UCI Machine Learning Repository: Irvine, CA, USA, 2010. [CrossRef]
17. Roggen, D.; Calatroni, A.; Rossi, M.; Holleczek, T.; Förster, K.; Tröster, G.; Lukowicz, P.; Bannach, D.; Pirkl, G.; Ferscha, A.; et al. Collecting complex activity datasets in highly rich networked sensor environments. In Proceedings of the 7th International Conference on Networked Sensing Systems (INSS), Kassel, Germany, 15–18 June 2010; pp. 233–240. [CrossRef]
18. Khan, M.; Hossni, Y. A comparative analysis of LSTM models aided with attention and squeeze and excitation blocks for activity recognition. *Sci. Rep.* **2025**, *15*, 3858. [CrossRef] [PubMed]
19. Lamaakal, I.; Essahraui, S.; Maleh, Y.; Makkaoui, K.E.; Ouahbi, I.; Bouami, M.F.; El-Latif, A.A.A.; Almousa, M.; Peng, J.; Niyato, D. A comprehensive survey on tiny machine learning for human behavior analysis. *IEEE Internet Things J.* **2025**, *12*, 32419–32443. [CrossRef]
20. Wu, J.; Ren, P.; Song, B.; Zhang, R.; Zhao, C.; Zhang, X. Data glove-based gesture recognition using CNN-BiLSTM model with attention mechanism. *PLoS ONE* **2023**, *18*, e0294174. [CrossRef] [PubMed]
21. Faisal, M.A.A.; Abir, F.F.; Ahmed, M.U.; Ahad, M.A.R. Exploiting domain transformation and deep learning for hand gesture recognition using a low-cost dataglove. *Sci. Rep.* **2022**, *12*, 21446. [CrossRef] [PubMed]
22. Shiranthika, C.; Premakumara, N.; Chiu, H.-L.; Samani, H.; Shyalika, C.; Yang, C.-Y. Human activity recognition using CNN and LSTM. In Proceedings of the 5th International Conference on Information Technology Research (ICITR), Moratuwa, Sri Lanka, 2–4 December 2020; pp. 1–6. [CrossRef]
23. Almujally, N.A.; Khan, D.; Al Mudawi, N.; Alonazi, M.; Alhasson, H.F.; Jalal, A.; Liu, H. Wearable sensors-based assistive technologies for patient health monitoring. *Front. Bioeng. Biotechnol.* **2025**, *13*, 1437877. [CrossRef] [PubMed]
24. González Garcés, M.; Cárdenas Montoya, J.; Peña Martínez, M.I.; Valencia García, J.; Hernández Rincón, E.H. Deep learning and noninvasive sensors for detecting physiological dysregulation: A scoping review. *J. Med. Syst.* **2026**, *50*, 14. [CrossRef] [PubMed]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.