Field Evaluation of a Wearable Framework for Worker Motion Recognition and Zone-Level Localization*

Minjun Chang[1], Sai Machiraju[2], and Yong K. Cho[3,†]

[1]School of Civil and Environmental Engineering, Georgia Institute of Technology, United States of America
[2]School of Electrical and Computer Engineering, Georgia Institute of Technology, United States of America
[3]Professor, School of Civil and Environmental Engineering, Georgia Institute of Technology, United States of America

minjunchang@gatech.edu, saimachi@gatech.edu,yong.cho@ce.gatech.edu

Abstract -

Unobserved worker motions on active construction jobsites contribute to injuries, near-misses, and productivity loss; however, continuous monitoring remains difficult under real deployment constraints. This study presents a real-time worker activity monitoring framework that integrates lightweight Temporal Convolutional Networks (TCN) with wearable sensing and zone-level localization to enable continuous and deployable behavior monitoring. The proposed system employs a three-IMU configuration mounted on the wrist, neck, and back to capture complementary upper-body motion while maintaining compatibility with personal protective equipment. Each sensor streams six-axis inertial data at 50 Hz and performs sliding-window inference using a TCN architecture with dilated causal one-dimensional convolutions and residual connections, supporting low-latency and memory-efficient processing suitable for edge deployment. Task-relevant construction activities are recognized using anatomically grounded motion patterns without relying on vision-based monitoring. To incorporate spatial context, Bluetooth Low Energy beacons provide RSSI-based zone localization that is temporally smoothed and synchronized with activity predictions to generate unified "what-and-where" behavioral labels. The framework is validated through a real factory deployment, demonstrating robust performance under occlusion, signal interference, and natural task variability. Results show that the system achieves reliable real-time inference with average accuracies of 76.28% at the motion level and 78.62% at the activity level, supporting scalable and privacy-preserving activity analytics in real construction environments. The field deployment further identifies key operational challenges, including sensor placement tradeoffs, signal occlusion effects, and user comfort constraints—and provides practical implementation guidelines for wearable monitoring systems in construction environments

Keywords -

Wearable Sensing; Worker Activity Recognition; BLE Localization; Edge AI; Construction Productivity

1 Introduction

Construction jobsites continue to experience injuries, near-misses, and productivity losses caused by unsafe or inefficient worker actions that often remain unobserved. Supervisors are typically responsible for monitoring multiple crews across dynamic work zones, making continuous observation difficult and limiting the ability to provide timely interventions. Although automated monitoring systems have been widely explored, many remain challenging to deploy in real construction environments due to occlusion, privacy concerns, infrastructure requirements, and operational disruptions.

Wearable sensing has emerged as a promising alternative for continuous worker monitoring because it is less sensitive to environmental visibility constraints and better supports privacy-preserving data collection. However, practical field deployment requires balancing recognition performance with usability, latency, and robustness. Many existing wearable monitoring approaches rely on dense sensor configurations or non-causal inference pipelines, which increase worker burden, computational requirements, and system complexity, thereby limiting adoption on active jobsites. To address these challenges, this study adopts a lightweight wearable configuration consisting of three sensing devices and applies a causal Temporal Convolutional Network (TCN) to enable low-latency motion recognition suitable for real-time edge deployment.

While motion recognition provides valuable behavioral information, it offers limited situational awareness without spatial context. Understanding where worker actions occur is critical for safety monitoring, task compliance evaluation, and productivity assessment. To provide this contextual information, the proposed framework integrates motion recognition with zone-level localization, enabling simultaneous interpretation of worker activities and spatial positioning.

This paper presents the first field-evaluated wearable

*This work was supported in part by the National Science Foundation (NSF) under Award No. 2515786 and by the Georgia Research Alliance (GRA). Any opinions, findings, conclusions, or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the NSF or GRA.

framework that tightly integrates causal TCN-based motion recognition with BLE zone-level localization into a unified real-time behavioral monitoring pipeline, validated on an active construction site with professional workers performing real installation tasks. The system is evaluated through laboratory testing and a real-world field case study to assess deployment feasibility and operational performance under realistic construction conditions. Rather than focusing on algorithmic development, this study emphasizes deployment-oriented evaluation through field implementation. Specifically, this case study investigates: (1) the minimum wearable sensor configuration that maintains monitoring performance while remaining acceptable to workers, (2) the reliability of activity recognition under uncontrolled field variability, including task transitions and signal noise, (3) deployment barriers that emerge during continuous real-time monitoring, and (4) stakeholder perceptions regarding the usefulness and usability of integrated activity and localization monitoring systems.

2 Literature Review

2.1 IMU based Human Motion Recognition

Wearable IMU-based motion recognition has been widely studied due to its robustness to visual occlusion and independence from environmental lighting conditions [1, 2]. High recognition accuracy has been reported in controlled laboratory settings, often using multiple IMUs distributed across the body [3, 4]. Kim and Cho [5] demonstrated automatic recognition of highway construction worker motions using wearable inertial sensors and Long Short-Term Memory (LSTM) networks, highlighting the effectiveness of deep sequence models for construction activity recognition.

Despite these advances, dense sensor configurations introduce challenges for field deployment, including increased setup time, user discomfort, interference with personal protective equipment, and maintenance overhead. To reduce sensing burden, prior studies have explored sparse sensor configurations or upper-body placements [3]. While these approaches improve usability, they often exhibit reduced robustness under natural task variability, motion transitions, and inter-subject differences. Most wearable sensing research remains laboratory validated, with limited reporting of long-duration field deployment outcomes.

In addition, many deep learning-based motion recognition methods rely on long temporal windows, non-causal inference, or offline post-processing [6]. These assumptions limit applicability for real-time field deployment, where bounded latency and causal operation are required. These gaps motivate the development of motion recognition systems designed explicitly for continuous on-site operation using minimal wearable instrumentation.

2.2 BLE-based Localization

Bluetooth Low Energy (BLE) beacon systems have been widely adopted for indoor localization due to their low infrastructure cost and deployment flexibility [7]. In construction and industrial environments, BLE-based approaches offer advantages where fixed camera systems or dense infrastructure are impractical. However, RSSI-based distance estimation is sensitive to multipath effects, body shadowing, and environmental interference [8], which are common on active jobsites.

As a result, many studies favor zone-level localization rather than high-precision tracking, using calibration and temporal filtering to improve robustness [7]. Zone-level localization aligns well with safety and productivity applications, where presence within a designated work area is often sufficient. Despite this, BLE localization is frequently evaluated independently of activity recognition systems, limiting its ability to provide actionable behavioral context. Integrating localization with motion recognition remains an open challenge for field-ready worker monitoring.

2.3 Construction Productivity

Construction productivity and safety research increasingly emphasizes the need for continuous and scalable monitoring systems [9]. Vision-based approaches provide rich contextual data but face persistent deployment barriers related to occlusion, privacy concerns, and maintenance in dynamic environments [10]. Wearable-based systems address these challenges but often focus on either motion recognition or localization rather than their integration.

Recent work highlights that actionable field monitoring requires joint understanding of worker actions and spatial context [11]. Knowing what motion is being performed and where it occurs enables zone-specific risk assessment, task compliance monitoring, and productivity analysis. However, many existing systems remain confined to laboratory validation or depend on complex sensing and computation that are difficult to sustain in the field. There remains a gap for field-ready frameworks that integrate lightweight motion recognition, practical localization, and real-time visualization while operating under realistic deployment constraints.

3 Methodology

The system is designed as two components: (1) IMU-based Motion Recognition, (2) BLE-based Localization. Each Personal Protection Unit (PPU) embeds one IMU and a Bluetooth signal receiver within a single microcontroller-based device, and workers are equipped with three PPUs to enable distributed motion and localization sensing.

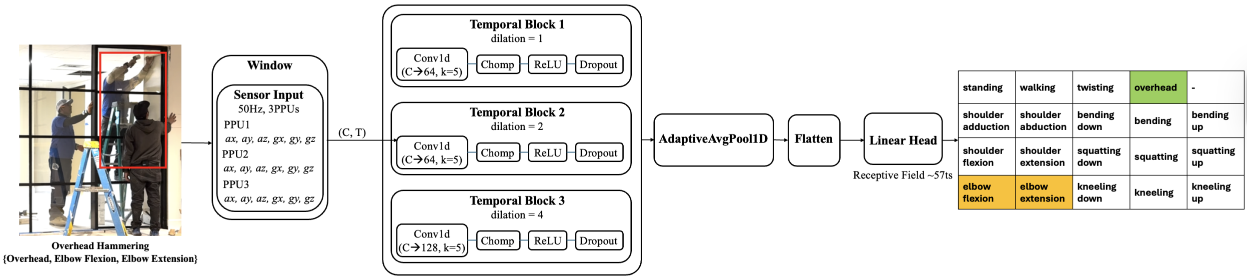

Figure 1. TCN Network Inference Pipeline

3.1 Motion Recognition

As illustrated in Figure 1, the motion recognition module performs real-time inference using a three-device wearable configuration. Three PPUs are mounted on the wrist, rear neck, and lower back to capture complementary distal and proximal body movements while remaining compatible with personal protective equipment. Each PPU integrates a 9-axis IMU consisting of an accelerometer, gyroscope, and magnetometer. Prior to deployment, the accelerometer and magnetometer are calibrated to correct hard-iron and soft-iron distortions by rotating the IMU along a spherical trajectory and fitting the sampled data to a sphere, yielding offset parameters and a correction matrix. The gyroscope is calibrated by averaging 500 stationary samples and subtracting the resulting bias from subsequent measurements. Although each IMU provides nine-axis data, only six-axis inertial measurements from the accelerometer and gyroscope are streamed at 50 Hz for inference, while magnetometer channels are excluded due to magnetic interference commonly encountered near steel structures and powered equipment on construction sites.

IMU streams are segmented into short sliding windows to bound latency and memory usage. Each window is processed independently using channel-wise z-normalization to reduce inter-subject variation and sensitivity to mounting differences. This design avoids reliance on long-term state, which can drift during extended field deployment, and preserves responsiveness during motion transitions.

A Temporal Convolutional Network (TCN) is used as the core recognition model. The network consists of three dilated residual blocks with 64, 64, and 128 channels, a kernel size of 5, and dilation rates of 1, 2, and 4. Input windows span 100 samples (approximately 2s at 50 Hz). All convolutions are causal, and predictions depend only on past samples within each window. No future frames, bidirectional context, or offline smoothing are used. Causality is enforced through right padding and truncation, ensuring consistent alignment between input windows and output predictions. This structure enables parallel computation and predictable inference time suitable for edge deployment.

The TCN outputs per-window motion probabilities, and a lightweight temporal smoothing step is applied over recent outputs to stabilize predictions during short transitions. This smoothing operates only on past predictions and introduces no look-ahead latency. The resulting pipeline achieves reliable real-time motion recognition while balancing accuracy, responsiveness, and deployability for field use.

3.2 Training Configuration

Model training and evaluation follow the protocol established in prior work to balance recognition performance and deployment practicality. IMU data are segmented into short sliding windows and processed using identical preprocessing across all experiments. Two training strategies are evaluated: a fixed single-split configuration with an 80:20 train/test split, and a leave-one-subject-out (LOSO) configuration for cross-subject generalization analysis. All models use the same Temporal Convolutional Network architecture, windowing parameters, and optimization settings so that observed differences reflect only sensor combinations and training strategies.

Seven sensor combinations are evaluated, as summarized in Table 1, including single-IMU, dual-IMU, and tri-IMU configurations. In-lab results indicate that multi-sensor configurations consistently outperform single-sensor setups. The wrist provides strong discriminative cues for upper-limb motion, while the neck and back contribute complementary proximal information. Among all tested configurations, the Neck+Back+Wrist (NBW) combination achieves the most stable performance across accuracy and F1 metrics in repeated in-lab inference trials. Although certain dual-sensor configurations achieve comparable peak accuracy, the NBW configuration demonstrates greater robustness during motion transitions and sustained operation.

Based on these results, the single-split (8:2) model trained with the NBW sensor combination is selected for the real-world case study described in Section 4. The single-split model is chosen over the LOSO model because

Table 1. Combinations and Lab Test Results

ID	Selected IMUs	Acc.	ID	Selected IMUs	Acc.
1	Neck+Back+Wrist (NBW)	0.8326	5	Neck (N)	0.7720
2	Neck+Back (NB)	0.7785	6	Back (B)	0.6878
3	Neck+Wrist (NW)	0.8353	7	Wrist (W)	0.7363
4	Back+Wrist (BW)	0.7926			

the case study evaluates continuous system performance under a fixed deployment setup rather than cross-subject generalization. This selection prioritizes temporal stability and real-time inference reliability, which are critical for evaluating system behavior during live monitoring in a factory environment.

3.3 Zone-Level Localization

Zone-level localization is implemented using commercial BC04P Bluetooth Low Energy (BLE) beacons deployed across the worksite. All beacons operate with identical advertising frequency and transmission intervals to maintain consistent signal behavior across zones. Transmit power is adjusted based on beacon placement height to balance coverage and signal stability. Beacons placed near floor level are configured at low transmit power (−16 dBm) to limit signal overlap and reduce multipath effects. Beacons mounted on ceilings higher than 10 ft are configured at higher transmit power, up to +4 dBm, to ensure adequate coverage across open work areas. The selected transmission settings also ensure long operational stability, with beacon battery life exceeding six months under the configured test conditions and extending beyond one year when operating at minimum transmit power.

Each wearable PPU continuously scans for BLE advertisements and records received signal strength indicator (RSSI) values from nearby beacons. Relative distances to beacons are estimated using a calibrated RSSI-to-distance model established during site configuration. To improve distance estimation accuracy, the PPUs apply an environment-specific signal propagation constant in the log-distance path loss model, which characterizes indoor attenuation conditions. This constant is determined experimentally by measuring RSSI values at multiple known distances from each beacon prior to deployment and typically ranges between 2 and 3 for indoor environments. The calibrated distance estimates are combined with known beacon coordinates defined on a preconfigured worksite map, and the PPU computes its instantaneous planar position (x_i, y_i) at each sampling step. Localization is performed at the zone level rather than high-precision tracking to reduce sensitivity to short-term RSSI fluctuations while remaining sufficient for safety and productivity analysis.

To improve robustness against signal noise, body shadowing, and transient interference, raw position estimates are temporally smoothed on the server side using a weighted averaging scheme. Let (x_i, y_i) denote the estimated planar position at time step i within a sliding window of length $N = 10$. Each estimate is assigned a confidence weight w_i derived from the corresponding BLE signal quality, such as normalized RSSI magnitude or beacon visibility count. The displayed position $(\bar{x}, \bar{y})$ is computed as:

$$\bar{x} = \frac{\sum_{i=1}^{N} w_i x_i}{\sum_{i=1}^{N} w_i}, \qquad \bar{y} = \frac{\sum_{i=1}^{N} w_i y_i}{\sum_{i=1}^{N} w_i} \tag{1}$$

This weighted averaging reduces the influence of low-confidence measurements while preserving responsiveness suitable for real-time visualization.

The resulting localization output is used to support higher-level site analytics. First, worker positions are compared against predefined work zone boundaries to determine whether a worker remains within assigned task areas or enters unauthorized zones. This enables detection of potential safety violations and unintended access. Second, temporal changes in estimated position are used to quantify travel distance and dwell time. By correlating movement distance with recognized motion labels, the system distinguishes between walking-intensive tasks and stationary work patterns. This joint interpretation of spatial displacement and motion recognition provides additional context for assessing task compliance and productivity without requiring fine-grained positioning accuracy.

4 Case Study

A case study was conducted at a warehouse construction site for window and door installation in Atlanta, Georgia, USA.

4.1 Data Collection for Feasibility Study

The case study was conducted at an indoor showroom construction site. At the time of deployment, major structural elements such as drywall partitions had already been installed, while interior construction activities were ongoing. Compared to large-scale building projects, the site covered a relatively small area with multiple workers operating in close proximity and dense wall structures, creating constrained movement paths and frequent line-of-sight interruptions. On the test date, workers were engaged in glass wall and door installation. Other tasks such as painting and silicone moulding were present, but the door installation team was selected because their work involved more dynamic and varied movements.

A floor map of the site was manually generated using laser-based distance measurements to define work zones and beacon locations. The environment included reinforced concrete walls and large metallic components that were frequently repositioned during installation, introducing potential signal obstruction and multipath interference. This setup provided a realistic and moderately challenging indoor environment for evaluating zone-level localization performance.

Figure 2. Workers wearing PPUs during work

4.2 Case Study Description

Three workers participated in the study, all of whom were professional contractors employed by Stellar W&D. Two participants served as test subjects for the case study evaluation, while data from the third participant were collected to support further refinement of the training model. All participants were right-handed and performed collaborative tasks related to glass wall and door installation. Each worker was equipped with three Personal Protection Units (PPUs) mounted on the rear neck, lower back, and right wrist, as illustrated in Figure 2. Each PPU integrates an ICM-20948 IMU. The wrist unit measures approximately 80mm in diameter and weighs under 100g including the battery, with a continuous operating life exceeding 4 hours at the configured 50 Hz streaming rate. Seven beacons were selected based on site geometry and empirical calibration to achieve full coverage of the installation work zone while minimizing signal overlap between adjacent zones.

The total study duration was 2 hours and 13 minutes. Seven Bluetooth Low Energy (BLE) beacons were deployed across the test site at floor level with transmit power 16dBm and advertising interval of 100ms, as shown in Figure 3, where blue markers indicate beacon locations. Four fixed cameras were installed to capture the full work process. These recordings were used solely for offline annotation and ground truth labeling of worker motions to support accuracy analysis and evaluation of the recognition model.

5 Result

5.1 Deployment Feasibility

Qualitative feedback was collected through interviews with a workers who wore the PPUs during the case study to assess wearability and potential interference with daily tasks. The participant reported that the PPUs mounted on the back and neck using clip attachments did not interfere with work activities. However, the wrist-mounted PPU was perceived as slightly bulky during tasks involving frequent wrist motion, such as hammering. The participant suggested that reducing the device size to approximately 50mm in diameter could further improve comfort and usability.

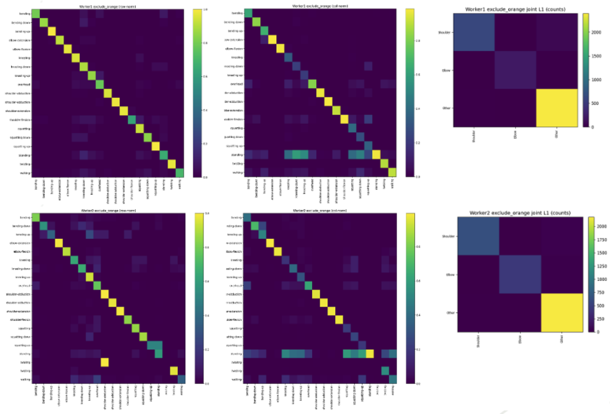

Figure 3. Worker 1 and 2 Confusion Matrices

In addition, the real-time worker localization and motion recognition interface was demonstrated to the site business owner to obtain feedback on practical value. The owner indicated that the combined presentation of worker location and motion provided meaningful insight for managing productivity and manpower allocation. In particular, the ability to observe where workers are located and what they are performing was identified as useful for task assignment and future planning. The owner also expressed interest in incorporating walking distance and task-specific activity distribution as additional indicators for more comprehensive productivity monitoring.

5.2 Feature Accuracy

Motion-level performance analysis was conducted using recognition data obtained from the pretrained Temporal Convolutional Network (TCN) model developed through in-lab data acquisition. The model was trained and validated under scripted laboratory scenarios, where in-lab inference tests achieved an average accuracy of 83.26%. This pretrained model was then deployed without modification during the case study to evaluate its feasibility and reliability under real field conditions.

Figure 3 presents confusion matrices for the case study results, with the top row corresponding to Worker 1 and the bottom row corresponding to Worker 2. The left column shows row-normalized confusion matrices, which reflect recall-oriented behavior by indicating how true motion classes are confused by the model. The middle column shows column-normalized confusion matrices, which emphasize precision by indicating how often predicted classes are correct. The right column presents joint-level accuracy, where shoulder- and elbow-related motions are concatenated to evaluate whether the model correctly identifies the active joint rather than the exact motion subtype. This analysis was performed to assess robustness under uncontrolled site conditions, which differ substantially from structured laboratory settings. The resulting motion-level accuracy shown in Table 2 was 75.86% for Worker 1 and 76.70% for Worker 2, demonstrating stable performance and supporting the feasibility of deploying the model on active construction sites. The lower F1 observed for Worker 1 is primarily attributed to the lab-to-field domain gap and a task composition dominated by overhead hammering sequences, which involve coupled elbow and shoulder motions that are historically the most confused class pair in training evaluation. Worker 2's higher F1 reflects a task distribution more consistent with the training set. Class-weighted training and domain adaptation are identified as direct future remedies.

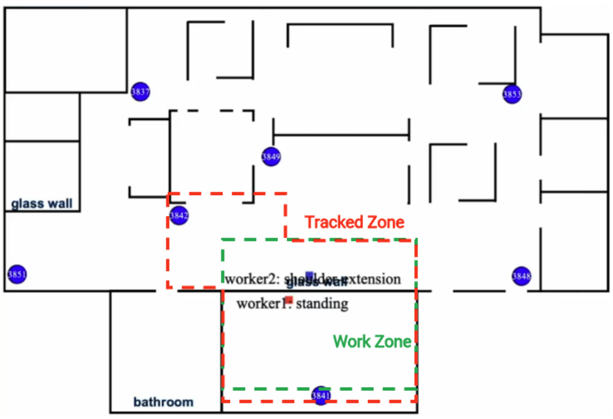

Figure 4. Work Zone and Tracked Zone Coverage

Activity-level accuracy was evaluated to examine the viability of the proposed method for assessing worker productivity in a more realistic setting. Activities performed during the test period were manually annotated, and each activity was defined as a sequence of constituent motion primitives, as illustrated in Figure 4. A recognized motion was considered correct if it corresponded to any motion primitive associated with the annotated activity at the given time frame. The resulting activity-level accuracy was 80.07% for Worker 1 and 77.16% for Worker 2, indicating that individual motion patterns influenced recognition performance while maintaining sufficiently high accuracy for practical productivity assessment.

Table 2. Motion/Activity Recognition Performance

Metric	Worker 1	Worker 2
Motion Level Accuracy	75.86%	76.70%
Motion F1-Score	0.5510	0.6401
Activity Level Accuracy	80.07%	77.16%

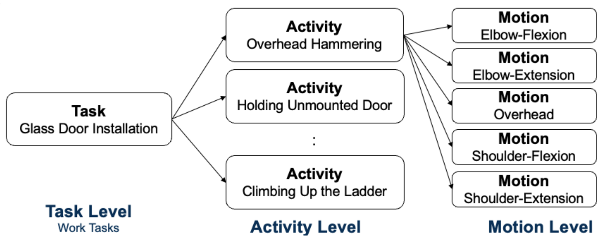

Figure 5. Task to Motion Breakdown

Localization performance was evaluated using area coverage as the primary metric. The designated work zone for the installation task covered approximately 774.8ft^2, while the area in which workers were visualized on the web-based interface covered approximately 563.5ft^2, as shown in Figure5 and Table3. Quantified localization error was not computed in this study as the primary objective was zone-level presence detection. The tracked area fully overlapped the defined work zone, resulting in 100% work zone coverage. However, approximately 27.27% of the tracked area corresponded to spatial estimates outside the intended zone. These inaccuracies are primarily attributed to variations in line-of-sight conditions between the wearable PPUs and nearby BLE beacons. For example, when a worker's body orientation obstructed direct signal paths, multiple PPUs occasionally received stronger RSSI measurements from an adjacent beacon, leading to temporary location estimates biased toward that beacon.

Table 3. Area Inspection Summary

Area Inspection	
Work Zone	≈ 774.8 ft^2
Tracked Zone	≈ 563.5 ft^2
Work Zone (WZ) Coverage	100%
WZ Portion	72.73%

6 Discussion

The case study evaluated the feasibility of deploying the proposed system under real field conditions and assessed its performance in real-time motion recognition and zone-level localization. Results indicate that the system can operate continuously in an active work environment using a minimal wearable configuration and lightweight localization infrastructure. Motion-level predictions remained stable throughout the deployment, and zone-level localization provided consistent spatial context suitable for safety monitoring and preliminary productivity assessment.

Activity-level analysis was conducted through manual annotation of motion sequences rather than automated inference. Individual activities were identified by aggregating consecutive motion-level predictions based on visual inspection and task context. While this approach enabled initial evaluation of task execution under real field conditions, it does not scale and introduces subjectivity; developing automated activity aggregation from motion primitives is therefore a direct priority for future work. Furthermore, productivity assessment requires distinguishing work-related actions from idle or unrelated behaviors, which is not explicitly addressed at the motion level. Developing automated activity aggregation and work-behavior classification methods is therefore essential for scalable productivity analytics.

Several deployment lessons emerged from the field study. The three-PPU configuration successfully captured overall body motion patterns with acceptable accuracy; however, detecting fine-grained arm movements without a dedicated forearm sensor remains challenging. In such cases, upper-limb activity must be inferred indirectly from torso inclination and whole-body kinematics, which may reduce sensitivity for localized hand tasks. Worker compliance also affects the interpretation of productivity metrics. Natural behaviors such as standing, conversing, or checking mobile devices occur intermittently during work hours and must be separated from task-related actions to avoid overestimation of productive time. In addition, the case study was limited to three workers within a single task team, and activity-level analysis was manually annotated. Although feasibility was demonstrated, scalability claims refer to the lightweight system architecture rather than the current validation scope. Future work will extend deployment to multiple task types, larger worker groups, and varied site conditions to substantiate broader applicability. The current evaluation is limited to two workers, which precludes formal statistical comparison across worker profiles or task types; future work will expand the cohort size.

Based on observations from the field deployment, several practical recommendations are identified to support future implementation of the wearable IoT system in construction environments:

- Sensor Placement: Wrist-mounted IMUs provide strong discriminative information for upper-limb motion but may introduce minor comfort concerns during repetitive hand-intensive tasks. Neck and lower-back placements offer stable proximal motion signals with minimal interference to worker movement and personal protective equipment. Multi-location sensing improves robustness under natural task variability and motion transitions.

- Beacon Deployment: Zone-level localization can be achieved using low-power BLE beacons strategically positioned to reduce signal overlap. Floor-level placement improves zone discrimination in partitioned indoor environments but may introduce location bias due to line-of-sight obstruction. Elevated placement increases coverage in open spaces but may require additional installation effort. Beacon spacing and transmit power should be adapted to site geometry and obstruction conditions to maintain localization reliability.

- Calibration and Field Deployment: RSSI calibration and IMU calibration should be conducted during initial system setup to capture site-specific signal propagation characteristics and improve data accuracy. Periodic verification of signal quality is recommended to account for environmental changes, material interference, and body-shadowing effects that may arise during active construction. In addition, standardized worker onboarding and device fitting procedures should be established to ensure consistent sensor placement, maintain data reliability, and reduce variability caused by improper attachment.

7 Conclusion

This paper presented a field-evaluated wearable framework that integrates real-time motion recognition and zone-level localization to support continuous worker monitoring in construction environments. By combining a lightweight Temporal Convolutional Network with a reduced IMU configuration and BLE-based localization, the system demonstrated stable performance under real deployment conditions without relying on cameras or dense sensing infrastructure. Results from a factory case study showed reliable motion-level and activity-level recognition, as well as consistent spatial coverage suitable for safety and productivity analysis. While activity-level inference and productivity assessment were conducted manually in this study, the findings highlight the feasibility of the proposed framework and motivate future work on automated activity aggregation and work-related behavior classification. Overall, the results suggest that the proposed approach provides a practical foundation for scalable, privacy-preserving worker monitoring on active construction sites.

References

[1] Francisco J. Ordóñez and Daniel Roggen. Deep convolutional and lstm recurrent neural networks for multimodal wearable activity recognition. *Sensors*, 16(1):115, 2016.

[2] Jianbo Yang, Minh Nhut Nguyen, Phyo Phyo San, Xiaoli Li, and Shonali Krishnaswamy. Physical activity recognition with mobile phones: Challenges, methods, and applications. *IEEE Signal Processing Magazine*, 32(4):52–59, 2015.

[3] Kinam Kim and Yong K. Cho. Effective inertial sensor quantity and locations on a body for deep learning-based worker motion recognition. *Automation in Construction*, 113:103126, 2020.

[4] Minsoo Park, Seongwoo Son, Yuntae Jeon, Dongyoung Ko, Mingeon Cho, and Seunghee Park. Scaffolding worker imu time-series dataset for deep learning-based construction site behavior recognition. *Advanced Engineering Informatics*, 65:103232, 2025.

[5] Kinam Kim and Yong K. Cho. Automatic recognition of workers' motions in highway construction by using motion sensors and long short-term memory networks. *Journal of Construction Engineering and Management*, 147, 2020.

[6] 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.

[7] Ramsey Faragher and Robert Harle. Location fingerprinting with bluetooth low energy beacons. *IEEE Journal on Selected Areas in Communications*, 33 (11):2418–2428, 2015.

[8] Faheem Zafari, Athanasios Gkelias, and Kin K. Leung. A survey of indoor localization systems and technologies. *IEEE Communications Surveys & Tutorials*, 21(3):2568–2599, 2019.

[9] Jochen Teizer, Brian S. Allread, Craig E. Fullerton, and Jimmie Hinze. Autonomous pro-active real-time construction worker and equipment operator proximity safety alert system. *Automation in Construction*, 19(5):630–640, 2010.

[10] Jian Gong and Carlos H. Caldas. Computer vision-based video interpretation model for automated productivity analysis of construction operations. *Journal of Computing in Civil Engineering*, 23(3):151–160, 2009.

[11] Arun Singh Chatterjee, Bivas Mitra, and Sandip Chakraborty. Accoustate: Auto-annotation of imu-generated activity signatures under smart infrastructure. *arXiv preprint arXiv:2112.06651*, 2021.