The Power of Periodicity: Exploiting Periodic UWB CIRs for Robust Activity Recognition with Attention-aware Multi-level Wavelet

Han Lin
Institute of Science Tokyo
Tokyo, Japan
lin@miubiq.cs.titech.ac.jp

Atsushi Nomura
Institute of Science Tokyo
Tokyo, Japan
nomura@miubiq.cs.titech.ac.jp

Kota Tsubouchi
LY Corporation
Tokyo, Japan
ktsubouc@lycorp.co.jp

Nobuhiko Nishio
Ritsumeikan University
Osaka, Japan
nishio@is.ritsumei.ac.jp

Masamichi Shimosaka
Institute of Science Tokyo
Tokyo, Japan
simosaka@miubiq.cs.titech.ac.jp

***Abstract*—In recent years, wireless sensing techniques, such as Wi-Fi and Ultra-Wideband (UWB) signals, have gained attention for activity recognition due to their ability to address privacy concerns associated with traditional computer vision methods. While UWB Channel Impulse Response (CIR) is believed to be prominent approach as well as Wi-Fi Channel State Information (CSI), research on its application in activity recognition remains limited. Previous studies have not fully explored the brevity of single measurements or the potential for feature extraction from CIRs. This paper presents a novel approach to robust device-free activity recognition by exploiting periodic UWB CIR samples. Utilizing multi-level wavelet packet decomposition (WPD) and a customized attention mechanism, the proposed method effectively combines multi-resolution features, improving recognition accuracy and reducing the need for extensive fine-tuning. Experiments conducted in various scenarios validate the performance of the proposed approach, with ablation studies demonstrating the superiority of multi-resolution analysis over short-time Fourier transform (STFT) and highlighting the cost efficiency of the method.**

***Index Terms*—UWB, channel impulse response, wireless sensing, device-free, human activity recognition, wavelet packet decomposition**

I. INTRODUCTION

Human Activity Recognition (HAR) plays a pivotal role in the development of intelligent systems, with applications spanning from healthcare monitoring to smart home automation, such as blood pressure monitoring and fall detection [1], [2].

Traditionally, HAR methods have been divided into wearable sensor-based and vision-based approaches. While wearable sensors provide high precision, they often lack user comfort and convenience [3]. Vision-based methods, although effective, are hindered by environmental factors and significant privacy concerns [4].

Overcoming these limitations, wireless sensing techniques have emerged as a promising alternative, offering device-free solutions that eliminate the need for wearables and avoid the privacy issues of vision-based systems [5], [6]. Among these, RF-based methods, including Wi-Fi Channel State Information (CSI) combined with deep learning, have shown strong potential for HAR [7]–[9].

Recently, Ultra-Wideband (UWB) technology has gained attention due to its increasing use in smart home automation and indoor localization [10]. UWB transmits signals over wide bandwidths with minimal energy, allowing high precision without interfering with other transmissions. Its integration into consumer devices, like iPhones, further expands its potential applications.

While UWB Channel Impulse Response (CIR) has shown improved performance over Wi-Fi CSI for HAR [11], [12], prior studies often rely on brief single CIR measurements, which fail to capture the dynamic nature of activities or differentiate between similar motions. Existing CIR processing methods, such as Short-Time Fourier Transform (STFT), also require extensive manual tuning, limiting their practicality.

To address these challenges, this research proposes the Attention-aware Multi-level Wavelet (AMW) method, which uses Wavelet Packet Decomposition (WPD) and a custom attention mechanism to efficiently combine multi-resolution features. AMW reduces manual tuning and improves recognition accuracy, demonstrating robust performance and significant classification improvements in device-free HAR across diverse scenarios.

The contributions of this paper include the following points:

- We propose a device-free activity recognition approach by effectively leveraging periodic CIR samples.
- We exploit the optimal time-frequency analysis with multi-level wavelet and attention mechanism by designing a framework named AMW.
- We show the superiority of proposed approach over previous practices by conducting comparison experiments in several typical scenarios.

II. Related work

A. Device-free HAR based on channel information

Device-free HAR primarily utilizes Wi-Fi CSI and UWB CIR to extract channel information. Wi-Fi CSI has been extensively studied for passive activity recognition. Chen et al. proposed a neural network framework integrating bidirectional and attention mechanisms for effective CSI data processing [9]. Schäfer et al. introduced Nexmon, a tool that simplifies data collection across platforms, delivering robust HAR results [13]. However, the performance of these deep learning models can be hindered by individual variances and environmental dynamics, which introduce intrinsic noise into CSI data, ultimately impacting accuracy [10].

UWB technology, characterized by its high temporal resolution and broad bandwidth, has demonstrated significant potential for passive activity recognition. Sharma et al. showed UWB CIR's effectiveness in basic activity classification [11], while Bocus et al. highlighted UWB's advantages over Wi-Fi CSI in specific scenarios [12].

B. Feature extraction for periodic signal

To address the time-variant nature of wireless signals in device-free HAR, time-frequency analysis techniques such as STFT and Discrete Wavelet Transform (DWT) have been widely applied to generate spectrograms [2], [14]–[17].

Multi-scale wavelet techniques have been particularly effective in robust feature extraction. Fang et al. employed multi-level discrete wavelet transform for CSI-based indoor positioning, while Wang et al. utilized multi-resolution analysis with wavelet packet functions to process UWB signals, constructing energy decomposition trees for CNN training [18], [19].

Despite their utility, these methods often require extensive manual hyperparameter tuning and fail to fully exploit resolution features. This limitation can lead to misrecognition in HAR, as specific resolutions are better suited for representing certain activities. Improving the utilization of resolution features remains an open challenge.

C. Attention-based methods

Recently, attention-based models have gained attention for their ability to reduce reliance on extensive parameter tuning. Fei et al. demonstrated the effectiveness of Vision Transformers (ViT) for Wi-Fi CSI analysis [20], while ConTransEn, a hybrid model combining CNNs and Transformers, achieved competitive results [21].

Time-frequency analysis has been shown to be effective in the literature, although extensive tuning efforts are needed as aforementioned. On the other hand, the reliance on large datasets and the complexity of attention mechanisms pose challenges in resource-constrained environments. A promising direction lies in integrating traditional signal processing techniques with attention mechanisms to address hyperparameter tuning issues and reduce training costs. Such a hybrid approach could leverage the strengths of both methodologies, potentially unlocking new capabilities for HAR in real-world applications.

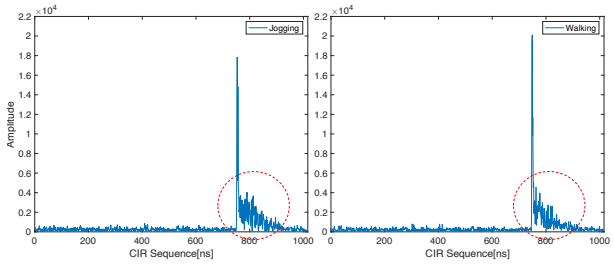


Fig. 1: Problem on differentiation of jogging and walking.

III. Problem Settings

A. Channel Impulse Response of UWB

The Channel Impulse Response of UWB represents encapsulates the channel characteristics, describing signals as a composition of multi-path components based on the Saleh-Valenzuela model [22]. The received signal $r(\cdot)$ of UWB system can be represented as follows:

$$r(t) = \sum_{k=0}^{K} \alpha_k e^{j\theta_k} s(t - \tau_k) + n(t) \tag{1}$$

Here, α_k, θ_k, and τ_k represent the amplitude, phase shift, and time delay of the k^{th} path, respectively. $s(\cdot)$ denotes the transmitted signal, and $n(\cdot)$ represents additive Gaussian noise. CIR effectively captures multi-path propagation, enabling detection of signal variations caused by human movement.

B. Limitations of Existing Researches

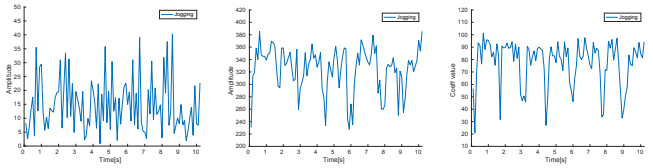
1) Brevity of Single CIR Measurement: Studies using single-shot CIR measurements for device-free HAR [11], [12] are limited by the brevity of CIR data (1[μs]), making it challenging to capture activities lasting several seconds. This hampers differentiation between similar activities like jogging and walking (Fig. 1).

To address this, multiple CIR samples over time can be used to capture complete activity durations. Although applied in other fields [23], [24], this approach remains underexplored in device-free HAR. This study leverages periodic CIRs to enhance activity recognition accuracy.

2) Feature Extraction from Periodic CIRs: Using shuttle jogging data, we explore feature extraction challenges. Fig. 2(a) shows that the raw time-domain amplitude of the first path struggles to depict activity variations due to significant noise in CIR data. Applying Fast Fourier Transform (FFT) removes noise by isolating relevant frequency ranges, as in prior studies [12], revealing periodic signal variations linked to activity movements (Fig. 2(b)).

To enhance localized analysis, third-level wavelet decomposition is performed to extract low-frequency components around the first path, effectively capturing the periodicity of jogging (Fig. 2(c)).

While time-frequency analysis shows promise, selecting optimal resolutions and reducing noise requires careful tuning of parameters (e.g., STFT window size, wavelet levels).



(a) Time series of raw first path amplitude (b) Time series of low frequency component (c) Time series of low frequency component around first path

Fig. 2: Time series features extracted from periodic CIRs

This process is computationally intensive and time-consuming, posing a challenge for practical implementations.

IV. Proposed Method

A. Overview

For the purpose of accurately capturing the full duration of activities, we utilize periodic UWB CIRs, comprising sequential CIR measurements over time. Our approach combines initial pre-processing with multi-level WPD to extract diverse time-frequency features. An attention mechanism is then applied to automatically identify and prioritize the most relevant resolutions, ensuring optimal feature representation with minimal manual parameter tuning. Fig. 3 illustrates the proposed method.

B. CIRs Acquisition and Pre-Processing

1) Periodic CIRs: In this study, we expand the measurements in the dimension of time-series to span the entire activity period, direct our attention to analyzing the amplitude of CIR. The periodic CIR matrix at time t, denoted by $\boldsymbol{M}^{(t)}$, is formulated as follows:

$$\boldsymbol{M}^{(t)} = \begin{bmatrix} \boldsymbol{x}^{(t-c\tau)} \\ \vdots \\ \boldsymbol{x}^{(t-\tau)} \\ \boldsymbol{x}^{(t)} \end{bmatrix} \tag{2}$$

where τ signifies the interval between successive CIR measurements(e.g., 0.1[s]). while c represents the count of measurements taken before the current measurement, respectively. The term $\boldsymbol{x}^{(t)}$ denotes the amplitude row vector of a single CIR measurement at time t, which is defined as:

$$\boldsymbol{x}^{(t)} = [|r(t)|, |r(t+\Delta T)|, \dots |r(t+N\Delta T)|] \tag{3}$$

with $N-1$ denoting the length of the single-shot CIR sequence, and ΔT representing the sampling interval of CIR measurements in the nanosecond range.

2) Calibration: As illustrated in Fig. 3, the raw CIR data exhibits redundancy along with significant noise, attributed to environmental factors and device output. To address this issue, we leverage the estimated first path index value according to the Leading Edge Detection algorithm [25] to discard noisy segments and focus on fluctuations caused by the multi-path effect. Consequently, the refined amplitude vector for a single CIR measurement is represented as:

$$\boldsymbol{x}'^{(t)} = [|r(t+(l_t-o_p)\Delta T)|, \dots |r(t+(l_t+o_f)\Delta T)|] \tag{4}$$

where the l_t indicates the first path index of CIR measurement at time t, o_p and o_f denote the offset around the index. The segment before the first path is retained to estimate noise levels.

Furthermore, we employ the Preamble Accumulation Count (PAC) value, which correlates with signal quality [26], mitigating environmental effects. The calibrated periodic CIR matrix is expressed as:

$$\boldsymbol{M}'^{(t)} = \begin{bmatrix} \boldsymbol{x}'^{(t-c\tau)}/p_{t-c\tau} \\ \vdots \\ \boldsymbol{x}'^{(t-\tau)}/p_{t-\tau} \\ \boldsymbol{x}'^{(t)}/p_t \end{bmatrix} \tag{5}$$

Here, the p_t denotes the PAC value associated with the CIR measurement at time t.

C. Multi-resolutions Feature Extraction

Unlike DWT, WPD allows for further decomposition of high-frequency components, where each node represents a set of coefficients corresponding to a specific frequency band [27], [28]. This makes WPD particularly suitable for analyzing UWB signals, as it captures the signal characteristics in the joint time-frequency domain by analyzing the coefficients of the resulting nodes [29].

The Haar wavelet was chosen for decomposition due to its optimal time-space resolution, robustness to frequency variations, and suitability for signals with abrupt transitions, such as those found in human activities [30].

D. Wavelet-Attentive Temporal Network(WATNet)

After extracting WPD coefficients as feature maps, these are input into a neural network for activity classification. An attention mechanism selectively integrates features across levels, while a Bidirectional Long Short-Term Memory (Bi-LSTM) network extracts time-series features from periodic CIRs. The network concludes with a flattening layer, a Fully Connected (FC) layer, and a softmax layer for classification, using CrossEntropyLoss as the loss function. The architecture is shown in Fig. 4.

To integrate features from multiple decomposition levels, we introduce a self-attention layer that operates on WPD coefficient dimensions, assigning weights across time-frequency resolutions. As depicted in Fig. 4, feature maps derived from multi-level WPD of calibrated periodic CIRs are processed by the attention mechanism. This mechanism computes scores for each coefficient and transforms them into attention-weighted maps. By learning the importance of different resolutions, this layer automates feature integration, assigning higher weights to the most informative features.

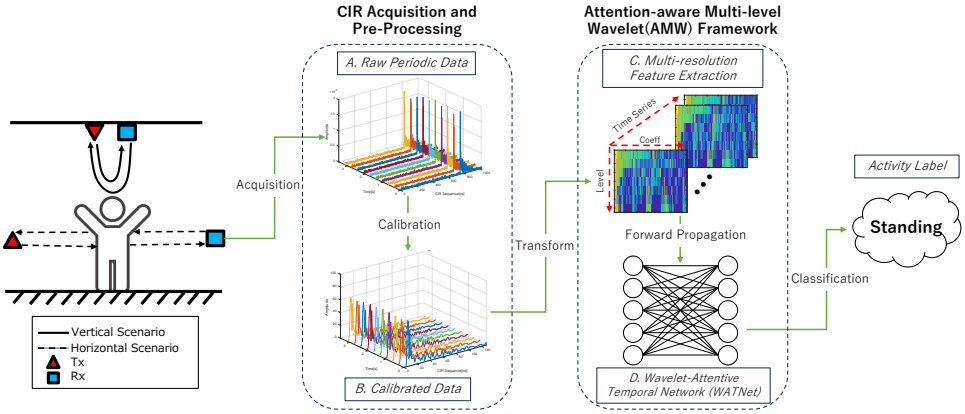


Fig. 3: Overview of proposal, consisting of CIRs Acquisition, Pre-Processing and AMW framework.

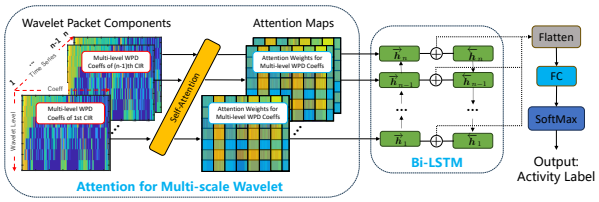


Fig. 4: The WATNet architecture for multi-level Wavelet Packet Components

V. Experiment Results

A. Experiment settings

1) Device Information: We conducted experiments using Decawave EVK1000 board pairs, implementing the Two-Way Ranging protocol for precise distance measurements per the 802.15.4a standard. To meet the specific configuration requirements of our study, we enabled simultaneous transmissions across multiple devices and set the maximum sampling rate to 10 Hz. The boards, equipped with DW1000 chips, operated at a 4.0 GHz carrier frequency, 500 MHz bandwidth, 64 MHz Pulse Repetition Frequency, 1024-bit preamble, and a 6.8 Mbps data rate. CIRs (1016 samples each) were extracted via laptops or Raspberry Pis.

2) Experiment Environments: The experiments were conducted in various settings, including a meeting room, office, corridor, and outdoor area, as illustrated in Fig. 5. While the meeting room served as a controlled environment, the other settings were designed to closely simulate natural conditions. To maintain the study's focus on single-person recognition, participants were required to cross the devices' transmission paths.

In the office environment, in addition to the participant, other individuals continued working around the sensor setup to simulate environmental noise caused by human activity. In the corridor, occasional passersby added to the complexity of the environment. The corridor's low ceiling and smaller space created conditions for simulating complex signal reflections from walls. The outdoor setting, though similar to the corridor in its experimental design, produced fewer reflections due to the absence of objects. Instead, it introduced environmental noise, such as wind, providing a unique challenge for evaluating system performance in outdoor scenarios.

Across these environments (excluding the outdoor setup), we tested two typical device placement scenarios:

i. **Horizontal scenario:** The implementation of horizontal device placement has been a prevalent methodology in prior researches, particularly in scenarios where device alignment at specific levels captures abundant human reflections effectively.

ii. **Vertical scenario:** Contrary to the horizontal arrangement, the vertical device arrangement, entailing device placement on the ceiling, significantly diminishes human reflections. The EVK1000 devices were encased in custom 3D-printed enclosures and securely mounted to the ceiling.

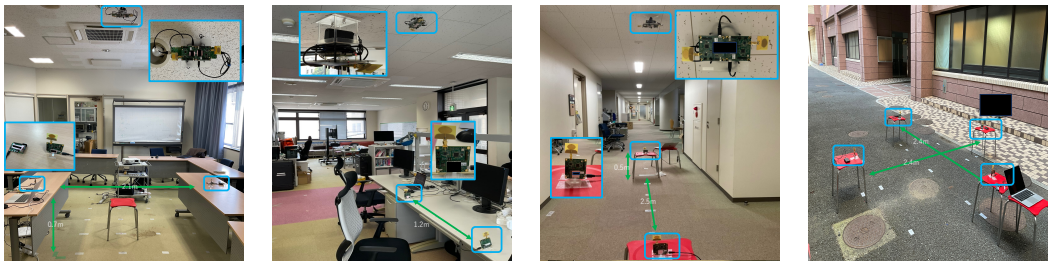
3) Datasets: We collected CIR data as participants performed specified activities within the environment involved six male adults. Our dataset encompasses four dynamic behaviors —jogging, walking, waving, and transitioning (between standing and sitting)—and two static postures: standing and sitting, as depicted in Fig. 6.

The waving activity involves subjects waving their right hands, and transitioning entails movements between standing and sitting. Furthermore, to validate human presence detection, a "No Activity" category was introduced.

4) Other Configurations: To account for the periodicity of each activity, as observed in the CIR variations shown in Fig. 2(c), the window size and stride are empirically set to 1.5[s] and 0.1[s], respectively, in the following experiments. F1 score is employed as a metric to evaluate activity recognition performance, utilizing a k-fold cross-validation approach where k is set to 10.

B. Comparison Experiment

1) Comparison With Previous Feature Extraction Methods: In this experiment, we compare the performance of existing feature extraction methods with our proposed approach. One



(a) Meeting Room (b) Office (c) Corridor (d) Outdoor

Fig. 5: Experiment environments

(a) Jogging (b) Walking (c) Waving (d) Transitioning (e) Standing (f) Sitting

Fig. 6: Activity classes for recognition including 4 dynamic activities and 2 static postures

TABLE I: Comparisons Results

Setting	**Peak-Value** [1] [23]	**Single-Shot** [11] [12]	**AMW (Ours)**
Meeting Room(H)	0.313	0.916	0.993
Meeting Room(V)	0.080	0.808	0.973
Office(H)	0.242	0.966	0.997
Office(V)	0.133	0.839	0.993
Corridor(H)	0.596	0.850	0.956
Corridor(V)	0.161	0.751	0.958
Outdoor	0.301	0.715	0.937

common method is to directly leverage single CIR measurements [11], [12]. As lack of similar studies on HAR with UWB CIR, the methods used in health monitoring is used to compare involving extracting signal changes at specific points and analyzing them as a sequence [1], [23].

Table. I presents the comparison results. The methods used in health monitoring show poor performance, as they fail to capture the detailed posture information represented by the entire CIR measurement. Additionally, activities are typically more dynamic and involve greater movement intensity compared to micro-motions like pulse monitoring, leading to more fluctuating CIR reflections.

While single-shot CIR can effectively identify static postures and detect human presence, it struggles to distinguish between dynamic activities such as Waving and Transitioning, and especially between Jogging and Walking due to their similar movement patterns but differing speeds. In contrast, our proposed method significantly improves recognition accuracy, particularly in scenarios with sparse human reflections, such as outdoor and vertical placements.

2) Comparsion with STFT: Since STFT is a widely used tool for time-frequency analysis, we fine-tuned its parameters to compare its performance against our proposed method. To streamline the parameter optimization process, the window function was fixed as a Hamming window, while the window length and overlap ratio (relative to the window length) were varied to correspond with each decomposition level of WPD. For a fair comparison, we use a standard Bi-LSTM combined with the STFT. The results across different environments in horizontal scenario are shown in Fig. 7.

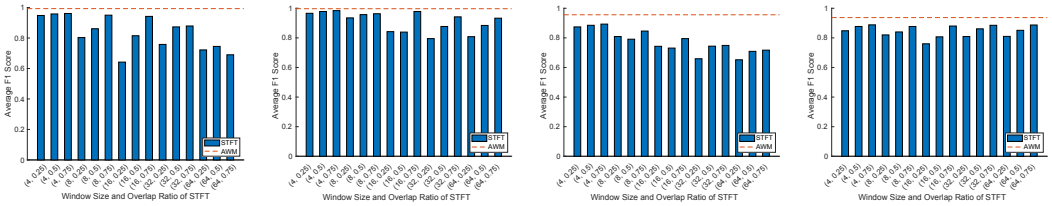
With parameter fine-tuning, the STFT can achieve performance comparable to our method. Although we have not exhaustively explored all possible parameter configurations for STFT, achieving optimal accuracy would require considerable manual effort and time. In contrast, our proposed method does not demand extensive fine-tuning to attain high performance, making it a more practical and efficient approach.

C. Ablation study on levels of Wavelet

This experiment evaluates WPD coefficients at various levels to compare time-frequency resolution performance for different activities. Three methods are examined: (1) Baseline: Utilizes periodic CIRs with only pre-processing; (2) Different Decomposition Levels: Applies specific levels of WPD to calibrated CIRs; (3) Multi-level WPD: Combines multiple levels, including “Best Levels” (Top 3 levels based on F1 scores), “MW” (without attention), and the proposed “AMW”. All methods, except AMW, use a vanilla Bi-LSTM model.

Taking the horizontal scenario results in the meeting room as an example, summarized in Table II, single-resolution features slightly outperform the baseline, with the fourth decomposition level achieving the best accuracy. However, finer levels, such as the seventh, reduce accuracy due to sparse features. Variability across activities highlights the importance of combining features from all levels for optimal performance.

The MW method improves accuracy by integrating all levels but struggles with similar activities like Jogging and Walking



(a) Meeting Room(H) (b) Office(H) (c) Corridor(H) (d) Outdoor

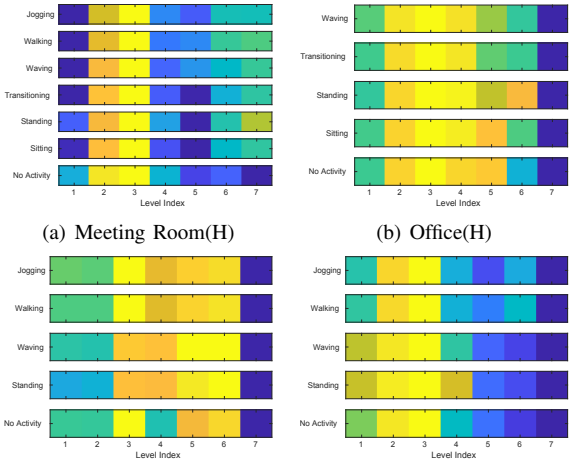
Fig. 7: Comparison between our method and STFT with fine-tuning.

TABLE II: 10-fold F1 score results of the Horizontal scenario in the Meeting Room

Method	Average	No Activity	Standing	Sitting	Waving	Walking	Jogging	Transitioning
Baseline	0.957	0.991	0.985	0.986	0.973	0.896	0.905	0.960
Level 1	0.964	0.993	0.985	0.983	0.980	0.914	0.925	0.965
Level 2	0.959	0.990	0.983	0.988	0.977	0.898	0.906	0.971
Level 3	0.965	0.991	0.982	0.990	0.979	0.917	0.926	0.972
Level 4	0.969	0.990	0.993	0.989	0.981	0.922	0.929	0.976
Level 5	0.967	0.993	0.992	0.991	0.981	0.911	0.923	0.975
Level 6	0.964	0.990	0.987	0.990	0.980	0.907	0.919	0.976
Level 7	0.956	0.989	0.985	0.985	0.975	0.887	0.899	0.968
Best Levels	0.977	0.996	0.991	0.991	0.988	0.943	0.950	0.982
MW	0.984	0.996	0.997	0.998	0.989	0.955	0.960	0.991
AMW	**0.993**	**0.999**	**0.999**	**0.999**	**0.997**	**0.978**	**0.981**	**0.997**

due to interference from less relevant levels. In contrast, our AMW method prioritizes useful levels, achieving a 99.3% F1 score—3.6% higher than the baseline—with near-perfect recognition of static postures and improved differentiation of dynamic activities.

To illustrate how WATNet assigns weights, attention maps were extracted and grouped by activity class for clarity. The generalized patterns, shown in Fig. 8, reveal level-specific attention distributions.



(a) Meeting Room(H) (b) Office(H)

(c) Corridor(H) (d) Outdoor

Fig. 8: Attention maps for all the activity classes.

The attention maps reveal distinct focus patterns across different levels for various activities, varying by environment due to differences in signal reflection and transmission paths. For instance, finer frequency resolution dominates in the corridor, while finer time resolution is prioritized in the office. Interestingly, attention patterns do not always align with single levels showing optimal performance in the ablation study. This suggests that the attention mechanism dynamically reweighs features to optimize overall performance, demonstrating an adaptive capability to diverse scenarios without prior knowledge of optimal resolution levels.

VI. Conclusion

This study tackles the limitations of device-free Human Activity Recognition (HAR) using Ultra-Wideband (UWB) Channel Impulse Response (CIR) by introducing a novel approach that combines periodic UWB CIRs with an Attention-aware Multi-level Wavelet (AMW) method. The AMW method applies multi-level Wavelet Packet Decomposition on CIR measurements, integrating time-frequency features and a custom attention mechanism to enhance recognition accuracy. Experimental results demonstrate the method's effectiveness, achieving an F1 score above 90% across various configurations and distinguishing between similar activities like jogging and walking, showcasing its robustness.

However, challenges remain for future work. The method lacks adaptability to new environments, necessitating cross-domain adaptation techniques to reduce retraining costs and improve environmental robustness. Additionally, the feasibility study is limited to simulated environments, restricting applicability to real-world scenarios. Future research should incorporate real-world experiments, considering obstacles and varying sensor setups, to validate the method's practicality and scalability for broader applications.

Acknowledgement

This work was partially supported by JSPS KAKENHI Grant Number 24K03015.

REFERENCES

[1] Z. Wang, B. Jin, F. Zhang, S. Li, and J. Ma, "UWB-enabled sensing for fast and effortless blood pressure monitoring," *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.*, vol. 8, no. 2, pp. 1–26, 2024.

[2] Y. Wang, K. Wu, and L. M. Ni, "Wifall: Device-free fall detection by wireless networks," *IEEE Trans. Mobile Computing*, vol. 16, no. 2, pp. 581–594, 2016.

[3] E. McAdams, A. Krupaviciute, C. Géhin, E. Grenier, B. Massot, A. Dittmar, P. Rubel, and J. Fayn, "Wearable sensor systems: The challenges," in *2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. IEEE, 2011, pp. 3648–3651.

[4] J. Wang, Y. Chen, S. Hao, X. Peng, and L. Hu, "Deep learning for sensor-based activity recognition: A survey," *Pattern recognition letters*, vol. 119, pp. 3–11, 2019.

[5] M. A. Al-qaness, "Device-free human micro-activity recognition method using wifi signals," *Geo-spatial Information Science*, vol. 22, no. 2, pp. 128–137, 2019.

[6] M. Pang, X. Yang, J. Liu, P. Li, F. Yan, and P. Chen, "Device-free activity recognition: A survey," in *China Conference on Wireless Sensor Networks*. Springer, 2020, pp. 223–243.

[7] D. Zhang, J. Ma, Q. Chen, and L. M. Ni, "An RF-based system for tracking transceiver-free objects," in *PerCom2007*. IEEE, 2007, pp. 135–144.

[8] S. Sigg, S. Shi, and Y. Ji, "RF-based device-free recognition of simultaneously conducted activities," in *Proc. the 2013 ACM conference on Pervasive and ubiquitous computing adjunct publication*, 2013, pp. 531–540.

[9] Z. Chen, L. Zhang, C. Jiang, Z. Cao, and W. Cui, "Wi-Fi CSI based passive human activity recognition using attention based BLSTM," *IEEE Trans. Mobile Computing*, vol. 18, no. 11, pp. 2714–2724, 2018.

[10] J. Yang, Y. Xu, H. Cao, H. Zou, and L. Xie, "Deep learning and transfer learning for device-free human activity recognition: A survey," *J. Automation and Intelligence*, vol. 1, no. 1, p. 100007, 2022.

[11] S. Sharma, H. Mohammadmoradi, M. Heydariaan, and O. Gnawali, "Device-Free activity recognition using Ultra-Wideband Radios," in *ICNC2019*, 2019, pp. 1029–1033.

[12] M. J. Bocus, K. Chetty, and R. J. Piechocki, "UWB and WiFi systems as passive opportunistic activity sensing radars," in *RadarConf21*, 2021, pp. 1–6.

[13] J. Schäfer, B. R. Barrsiwal, M. Kokhkharova, H. Adil, and J. Liebehenschel, "Human activity recognition using CSI information with Nexmon," *Applied Sciences*, vol. 11, no. 19, p. 8860, 2021.

[14] J. Ding and Y. Wang, "Wi-Fi CSI-based human activity recognition using deep recurrent neural network," *IEEE Access*, vol. 7, pp. 174 257–174 269, 2019.

[15] D. Kumar, A. Sarkar, S. R. Kerketta, and D. Ghosh, "Human activity classification based on breathing patterns using IR-UWB radar," in *INDICON*. IEEE, 2019, pp. 1–4.

[16] Y. Mei, T. Jiang, X. Ding, Y. Zhong, S. Zhang, and Y. Liu, "WiWave: WiFi-based human activity recognition using the wavelet integrated CNN," in *ICCC Workshops 2021*. IEEE, 2021, pp. 100–105.

[17] H. Li, X. He, X. Chen, Y. Fang, and Q. Fang, "Wi-motion: A robust human activity recognition using Wi-Fi signals," *IEEE Access*, vol. 7, pp. 153 287–153 299, 2019.

[18] S.-H. Fang, W.-H. Chang, Y. Tsao, H.-C. Shih, and C. Wang, "Channel state reconstruction using multilevel discrete wavelet transform for improved fingerprinting-based indoor localization," *IEEE Sens. J.*, vol. 16, no. 21, pp. 7784–7791, 2016.

[19] J. Wang, K. Yu, J. Bu, Y. Lin, and S. Han, "Multi-classification of UWB signal propagation channels based on one-dimensional wavelet packet analysis and CNN," *IEEE Trans. Vehicular Technology*, vol. 71, no. 8, pp. 8534–8547, 2022.

[20] F. Luo, S. Khan, B. Jiang, and K. Wu, "Vision transformers for human activity recognition using wifi channel state information," *IEEE Internet of Things Journal*, vol. 11, no. 17, pp. 28 111–28 122, 2024.

[21] F. Ge, Z. Yang, Z. Dai, L. Tan, J. Hu, J. Li, and H. Qiu, "Human activity recognition based on self-attention mechanism in wifi environment," *IEEE Access*, vol. 12, pp. 85 231–85 243, 2024.

[22] A. Saleh and R. Valenzuela, "A statistical model for indoor multipath propagation," *IEEE J. Selected Areas in Communications*, vol. 5, no. 2, pp. 128–137, 1987.

[23] S. Li, B. Jin, F. Zhang, Z. Wang, J. Ma, X. Ren, and H. Liu, "Hypnos: A contactless sleep stage monitoring system using UWB signals," *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.*, vol. 8, no. 3, pp. 1–27, 2024.

[24] M. De Sanctis, A. Conte, T. Rossi, S. Di Domenico, and E. Cianca, "CIR-based device-free people counting via UWB signals," *IEEE Sens. J.*, vol. 21, no. 9, p. 3296, 2021.

[25] I. Sharp, K. Yu, and Y. J. Guo, "Peak and leading edge detection for time-of-arrival estimation in band-limited positioning systems," *IET communications*, vol. 3, no. 10, pp. 1616–1627, 2009.

[26] M. J. Bocus and R. Piechocki, "A comprehensive ultra-wideband dataset for non-cooperative contextual sensing," *Scientific Data*, vol. 9, no. 1, p. 650, 2022.

[27] A. K. M. Baki and N. C. Karmakar, "Improved method of node and threshold selection in wavelet packet transform for UWB impulse radio signal denoising," *Progress In Electromagnetics Research C*, vol. 38, pp. 241–257, 2013.

[28] K. J. Friesen, N. Panagiotacopulos, and W. L. Sjogren, "Analysis of gravity signals and gravity potential determination using order 8 multiresolution analysis discrete wavelets," in *Proc. Adv. Phys. Electron. Signal Process. Appl. WSES Press, Athens*. Citeseer, 2000, pp. 303–308.

[29] J. Li, X. rong Cui, H. Zhang, and T. A. Gulliver, "An UWB ranging method based on wavelet packet decomposition," *Neurocomputing*, vol. 270, pp. 75–81, 2017, distributed Control and Optimization with Resource-Constrained Networked Systems.

[30] W. Xi, D. Huang, K. Zhao, Y. Yan, Y. Cai, R. Ma, and D. Chen, "Device-free human activity recognition using CSI," in *Proc. the 1st Workshop on Context Sensing and Activity Recognition*, 2015, pp. 31–36.