



TagSleep3D: RF-based 3D Sleep Posture Skeleton Recognition

CHEN LIU, Northwest University, China and Xi'an Key Laboratory of Advanced Computing and System Security, China

ZIXUAN DONG, LI HUANG, WENLONG YAN, and **XIN WANG**, Northwest University, China

DINGYI FANG, Northwest University, China and Shaanxi International Joint Research Centre for the Battery-Free Internet of Things, China

XIAOJIANG CHEN, Northwest University, China and Xi'an Advanced Battery-Free Sensing and Computing Technology International Science and Technology Cooperation Base, China

Sleep posture plays a crucial role in maintaining good morpheus quality and overall health. As a result, long-term monitoring of 3D sleep postures is significant for sleep analysis and chronic disease prevention. To recognize sleep postures, traditional methods either use cameras to record image data or require the user to wear wearable devices or sleep on pressure mattresses. However, these methods could raise privacy concerns and cause discomfort during sleep. Accordingly, the RF (Radio Frequency) based method has emerged as a promising alternative. Despite most of these methods achieving high precision in classifying sleep postures, they struggle to retrieve 3D sleep postures due to difficulties in capturing 3D positions of static body joints. In this work, we propose TagSleep3D to resolve all the above issues. Specifically, inspired by the concept of RFID tag sheets, we explore the possibility of recognizing 3D sleep posture by deploying an RFID tag array under the bedsheets. When a user sleeps in bed, the signals of some tags could be blocked or reflected by the sleep posture, which can produce a body imprint. We then propose a novel deep learning model composed of the attention mechanism, convolutional neural network, and together with two data augmentation methods to retrieve the 3D sleep postures by analyzing these body imprints. We evaluate TagSleep3D with 43 users and we totally collect 27,300 sleep posture samples. Our extensive experiments demonstrate that TagSleep3D can recognize each joint on the human skeleton with a median MPJPE (Mean Per Joint Position Error) of 4.76 cm for seen users and 7.58 cm for unseen users.

CCS Concepts: • Human-centered computing → Ubiquitous and mobile computing systems and tools.

Additional Key Words and Phrases: Sleep Postures, 3D Skeletons, RFIDs, Neural Network

ACM Reference Format:

Chen Liu, Zixuan Dong, Li Huang, Wenlong Yan, Xin Wang, Dingyi Fang, and Xiaojiang Chen. 2024. TagSleep3D: RF-based 3D Sleep Posture Skeleton Recognition. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 8, 1, Article 13 (March 2024), 28 pages. <https://doi.org/10.1145/3643512>

The corresponding author: Dingyi Fang.

Authors' addresses: **Chen Liu**, Northwest University, Xi'an, China and Xi'an Key Laboratory of Advanced Computing and System Security, China; **Zixuan Dong**; **Li Huang**; **Wenlong Yan**; **Xin Wang**, Northwest University, Xi'an, China; **Dingyi Fang**, Northwest University, Xi'an, China and Shaanxi International Joint Research Centre for the Battery-Free Internet of Things, Xi'an, China; **Xiaojiang Chen**, Northwest University, Xi'an, China and Xi'an Advanced Battery-Free Sensing and Computing Technology International Science and Technology Cooperation Base, China.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

© 2024 Copyright held by the owner/author(s). Publication rights licensed to ACM.

ACM 2474-9567/2024/3-ART13

<https://doi.org/10.1145/3643512>

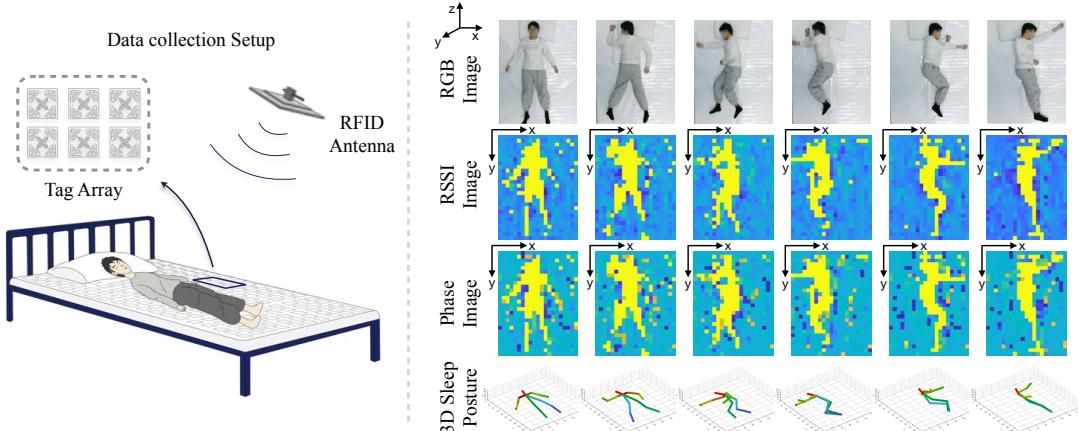


Fig. 1. Left: A non-wearable and in-home 3D sleep posture recognition system using RF signals. Right: The reconstructed 3D sleep pose skeleton from the captured RSS and phase of the RFID tags.

1 INTRODUCTION

Sleep posture can have a major impact on sleep quality, sleep disorder, overall health, and even personalities [13, 19, 78]. Past studies revealed that sleep posture is a crucial factor in alleviating or diagnosing various chronic diseases [12, 37, 40, 83, 87], such as sleep apnea [61, 88, 107], snoring [79], insomnia [26], sudden infant death syndrome (SIDS) [63], and pressure ulcers [24, 27, 65]. However, most existing sleep posture recognition systems focused on posture classification [50, 51], e.g., supine, prone, and lateral, which lacks detailed information on body and limb positions. Augmenting the sleep posture monitoring with the capability to sense the 3D sleep posture skeleton is crucial for achieving comfortable sleep [68] and diagnosing diseases [88, 107]. For instance, placing legs in an awkward position may cause leg cramps, while sleeping with hands up position adds continual pressure on the back causing shoulder pains. Moreover, the specific limb poses and body positions are important indicators for monitoring medical conditions, such as restless legs syndromes (RLS), periodic limb movement disorder (PLMD) [68], and Parkinson’s disease [16, 31, 86]. In medical settings, 3D sleep posture sensing aids physical rehabilitation [3, 17], orthopedic/spinal care, and enhances smart bedding systems to adjust firmness, positioning, or temperature for optimal comfort and support [28, 67, 74, 77]. Therefore, the capability to sense 3D sleep postures represents a significant advancement in sleep monitoring and healthcare.

Typically, sleep postures can be recognized through vision-based methods [18, 32, 48, 64, 97, 100] and wearable sensor-based methods, such as the chest-worn device [84, 103] and smartwatch [11]. However, these methods may cause either privacy concerns or discomfort during sleep, especially for the elderly [23]. More importantly, current wearable sensor-based methods can only estimate the sleep posture types, but cannot provide the 3D sleep posture skeleton. Although some works employ on-bed sensors (e.g., the pressure bedsheets) to estimate the 3D skeleton and even 3D body mesh of sleep postures from the pressure images [10, 14], the on-bed sensors may also affect sleep comfort and the deployment cost of the high-resolution sensor array is relatively high [51]. Hence, most of these approaches are only evaluated in labs. For non-wearable and in-home sleep posture monitoring, some recent works deploy Wi-Fi, RFID, FMCW, or mmWave radar near the bed [2, 50, 52, 56, 57, 102] or RFID tags under a bed sheet [51] to recognize the sleep posture by analyzing the changes of wireless signal features. However, most RF-based systems only focus on recognizing sleep posture types or body orientations. Although Argosleep [2] can retrieve the 3D sleep posture skeleton using mmWave radar, it requires an extra height classifier to estimate the user’s height to ensure a high recognition accuracy.

In this paper, we propose TagSleep3D, aiming to achieve robust 3D sleep posture skeleton recognition by using RF signals. Intuitively, we would apply existing RF-based localization methods to locate the keypoints of the human skeleton. However, the RF signals are vulnerable to multipath effects, which will lead to estimation errors. Another possible direction is to apply existing RF-based human 3D skeleton estimation approaches [36, 104, 105] which rely on dynamic motions, such as people moving and walking, to reconstruct the skeleton. However, when a person falls asleep in bed, the person's body is mostly static. Hence, these approaches are not proper for sleep posture recognition. To avoid the above problems, inspired by the concept of RFID tag sheets [51], we explore the possibility of recognizing 3D sleep posture skeletons by deploying an RFID tag array under the bedsheets. As shown in Fig. 1 Left, the antenna is placed on the ceiling right above the bed, while a number of tags are composed of a tag array. The antenna is powered by an RFID reader, and the reader receives the backscattered signals from the RFID tags. When the person sleeps in bed, the backscattered signals of some tags will be blocked or reflected by the person's body, which can produce an imprint of the human body as shown in Fig. 1 Right. Therefore, instead of locating the body's keypoints or relying on dynamic motions, we can reconstruct the 3D sleep posture skeleton by analyzing the imprints (i.e., RSSI/phase images denoted in Section 3.1) using the deep learning technique. However, it is non-trivial to realize TagSleep3D due to the following **challenges**:

- *How to recover 3D information from the incomplete 2D RSSI/phase images?* Although the RSSI/phase image depicts an imprint of the human body in the 2D plane, the information on the third dimension is missing, especially when some body parts do not have notable physical contact with the bedsheets.
- *How to achieve high recognition accuracy with low-resolution body imprints?* Due to the low resolution of RSSI/phase images, different sleep postures, particularly different limb poses, may result in similar imprints. Hence, it is challenging to train an accurate model in a less rich domain.
- *How to achieve robust performances across users and environmental setups?* The sleep posture varies across users and the RF signals are sensitive to environment changes. The tag array may also be affected by the mattress hardness, the bed size, and the antenna placement. All the above factors may cause unreliable results. However, it is time-consuming to collect a large number of training samples with rich sleep postures under various settings. Therefore, how to deal with the diversity of both sleep postures and environmental settings when training the model is challenging.

To solve these challenges, first, we reveal that the blocked tags and the unblocked tags in the tag array are complementary for sensing the 3D sleep posture skeleton. On one hand, the blocked tags provide a clear imprint of the human body in the 2D plane. On the other hand, the signal of the unblocked tags could be easily affected by different body parts through multi-path reflections, which can provide the third dimension information of the body parts in the 3D space. Second, we design a novel deep learning model incorporating the attention mechanism into the Convolutional Neural Network (CNN) to effectively utilize the features from both RSSI and phase, enabling an accurate 3D sleep posture skeleton recognition. Last, we propose two data augmentation methods to enrich the diversity of sleep postures and bolster data variability across environmental settings, ultimately enhancing the robustness of TagSleep3D.

Contributions: The main contributions of TagSleep3D are summarized as follows:

- TagSleep3D is the first work that can achieve robust 3D sleep posture skeleton recognition by using RFID tags, which has opened up new possibilities for using RF signals to sense static 3D pose.
- We reveal the complementarity between the blocked tags and unblocked tags in regard to 3D sleep postures. Based on this, we propose two data augmentation methods and a novel deep-learning model composed of the attention mechanism and CNN, achieving robust 3D sleep posture recognition across users and environments.
- We evaluate the performance of TagSleep3D with extensive experiments involving 43 users aged from 19 to 60. We have totally collected 27,300 sleep posture samples. The experimental results reveal that TagSleep3D

can recognize each joint on the human skeleton with a median MPJPE of 4.76 cm for seen users and 7.58 cm for unseen users. The average errors along the x, y, and z axes are 3.72 cm, 4.43 cm, and 3.45 cm, respectively for unseen users.

2 RELATED WORK

In recent years, there has been a growing interest in sleep posture recognition for sleep monitoring and healthcare. These works can be broadly classified into the following categories.

2.1 Camera-based Approaches

Existing camera-based approaches leverage color [30], infrared [18, 46, 64], thermal [32], depth cameras [6, 25, 48, 60, 101], or a combination of them [30, 47, 55, 85] to recognize sleep postures. Most of these works focused on posture classification, such as supine, prone, and lateral postures. Although some works can capture a 3D image of sleep by using depth cameras [39, 62, 71] or multi-modal sensor cameras [8, 97, 100], the camera-based approaches still suffer from inherent disadvantages including privacy concerns, lighting conditions, and blanket occlusion [49].

2.2 Wearable Sensor-based Approaches

Laboratory-based Polysomnography (PSG) is a professional sleep monitoring approach used in clinics [45]. It requires the patient to wear a dozen of dedicated devices/sensors, which are expensive and uncomfortable for long-term monitoring. For in-home sleep posture monitoring, wearable sensor-based solutions, such as the chest-worn device (e.g., chest belt [84, 103]), smartwatch (e.g., SleepGuard [11] and Sleepmonitor [81]), wristband (e.g., FitBit [21], Jeon et al. [35], Jawbone [33], and Jeng et al. [34]), and even sensor-augmented sleepwear (e.g., Phyjama [41, 42]) are designed to track the movement of the limbs based on the gravity sensor, accelerator, gyroscope, magnetometer, inertial sensor, and textile sensor readings. However, these works can only recognize sleep posture types and cannot provide the 3D pose skeleton. Moreover, wearing sensors/devices may cause discomfort to the user, especially for the elderly [23].

Furthermore, some approaches employ wearable RFID (e.g., RF-Kinect [90]) or integrate the information from high-speed cameras and wearable sensors (e.g., Vicon [89]) to retrieve the 3D body skeleton. However, these approaches are designed for body movement tracking, which is either expensive or not convenient for using in sleep scenarios due to the uncomfortable wearing experience and lighting conditions.

2.3 Pressure Bedsheet-based Approaches

A pressure-sensitive bedsheet is designed with an array of sensors, such as pressure sensors, accelerometers, fluid-filled cells, or a force-sensing array. The pressure bedsheet-based approaches are able to recognize sleep postures by classifying the captured pressure images [29, 53, 54, 70, 82, 95, 96, 106]. More recent works can estimate the 3D skeleton and even 3D body mesh of sleep postures from the pressure images by applying deep learning techniques [10, 14]. However, the on-bed sensors may affect sleep comfort and the pressure bedsheet is sensitive to self-occlusion. Further, the deployment cost of the pressure bedsheet is high due to its high-resolution pressure-sensor array and the power supply. As a result, most of these approaches are evaluated in labs and are not popular for in-home use. Similar to the pressure bedsheet-based approaches, our TagSleep3D employs a commercial RFID tag array under the bedsheet to recognize the 3D skeleton of sleep postures. Different from the high-resolution pressure-sensor array, RFID tags are cheap and comfortable¹, and do not require the wired power supply, which is promising for in-home monitoring.

¹A RFID tag is just like a piece of paper, and it would not change the feel of the bed.

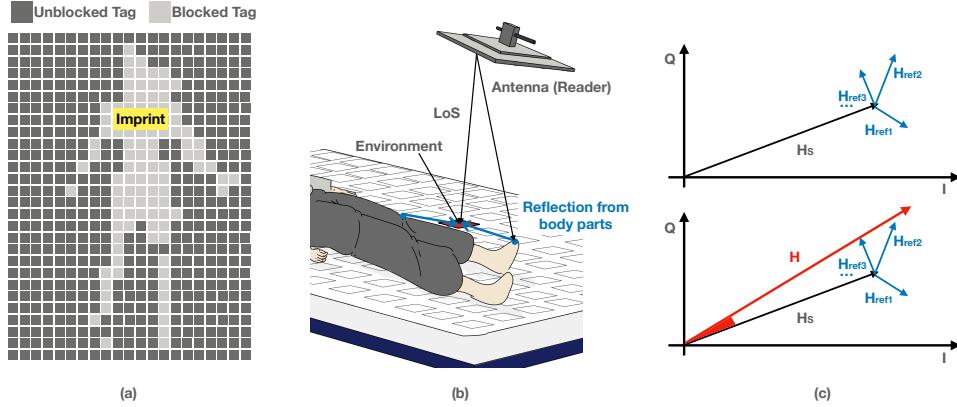


Fig. 2. (a) The blocked tags and unblocked tags demonstrate a clear sleep posture imprint. (b) The RFID signal propagation paths when a person sleeps in bed. (c) The representation of the superposition signal in the IQ domain.

2.4 RF-based Approaches

RF signals, such as Wi-Fi, RFID, and FMCW radio, have been proposed to monitor sleep postures in a non-contact manner [2, 5, 44, 50, 52, 56, 57, 76, 99, 102]. These approaches deploy the RF equipment near the bed and identify the specific sleep postures by analyzing the changes of wireless signal features including the RSSI (received signal strength indicator) [5], the CSI (channel state information) [52, 56, 57], the phase [50], and the reflections of RF signals [102]. Besides, some works deploy an array of RFID tags under a bed sheet [51] to detect sleep postures. However, most of the above systems can only recognize the sleep posture types or body orientation. Although Argosleep [2] can retrieve the 3D sleep posture skeleton using mmWave radar, it requires a height classifier to estimate the user’s height before predicting the joint locations. TagSleep3D differs from the existing systems by utilizing deep neural networks to recognize not only sleep posture types but also detailed 3D body skeletons. Our system, bolstered by a diverse dataset and advanced data augmentation techniques, ensures robust performance across different users and environments.

Note that past works also demonstrate the capability of estimating human 3D skeletons using FMCW radio [104, 105], mmWave [1, 20, 92], and Wi-Fi [36, 72, 73]. However, these approaches rely on dynamic motions, such as people moving and walking, to reconstruct the skeleton by combining different body parts across different RF snapshots [105] or by estimating the rotation of the joints [36]. In contrast, when a person falls asleep in bed, the person’s body is mostly static. Hence, the above approaches that rely on dynamic motions are not proper for sleep posture recognition.

2.5 Tag-array based Sensing

The passive RFID tag array has been widely used to sense the context of human targets, such as location tracking (e.g., Tadar [98]), vital signs monitoring (e.g., RF-ECG [91] and Badge [66]), activity recognition (e.g., RFitness [80]), and facial recognition (e.g., RFaceID [58], RFace [94]). These systems typically utilize small tag arrays, consisting of dozens of tags, to achieve high performance for their specific sensing objectives. Unlike these works, our work employs a larger tag array, comprising hundreds of tags, to sense sleep postures. This larger array can provide richer information and we employ deep neural networks in our system for fine-grained 3D sleep posture sensing.

3 PRELIMINARY

In this section, we introduce the background knowledge and the basis of our system.

3.1 The Representation of RF Signals

For 3D sleep posture skeleton recognition, we only need to process the static periods. Therefore, we remove the dynamic periods such as turning over by a motion detector [50]. For each static period, we employ a time window as a unit for signal segments, typically set at 3 seconds [51]. Within each time window t_i , we calculate the mean RSSI and phase measurements for each tag. Considering a $M \times N$ tag array, its RSSI and phase for time window t_i can be represented as an RSSI matrix $R_i = \{r_{m,n}^i\}$ and a phase matrix $P_i = \{p_{m,n}^i\}$ ($0 < m \leq M, 0 < n \leq N$). To mitigate the tag diversity, we adopt the RSSI (phase) difference between someone lying in bed and an empty bed for sleep posture recognition. Specifically, when no one is sleeping on the bed, we collect RF signals of the tag array for a few seconds and use the mean RSSI and phase measurements during this period for each tag. Accordingly, we have baseline matrices R_{empty} and P_{empty} . Then, we employ the signal difference matrices $\Delta R_i = |R_i - R_{\text{empty}}|$ and $\Delta P_i = |P_i - P_{\text{empty}}|$ to recognize the sleep posture in the time window t_i .

RF images. For clarity, we define ΔR_i and ΔP_i as an *RSSI image* and a *phase image* for the time window t_i . These two images are denoted as RF images. Note that, for a new setup, we only need to collect R_{empty} and P_{empty} once. Fig. 1 (Right) shows several examples of RSSI images, phase images, and their corresponding sleep postures. As we can see, the RSSI image and phase image can produce body imprints of the sleep posture.

3.2 Basis of TagSleep3D

When a user lies in the tag-array bed, the presence of the body between tags and the antenna will affect the backscattered signals of the tag array. On one hand, the signals of some tags will be blocked by the body, especially for those tags that are under the body, resulting in an imprint of the human body as shown in Fig. 2(a). Thus, this body imprint can be considered as a *unique feature* for sleep posture recognition. On the other hand, for the tags that are not blocked, there are additional reflection paths from different body parts as shown in Fig. 2(b). Hence, the signals of the unblocked tag are the superposition of the LoS signal and reflection signals from body parts and the environment, which can be expressed as

$$H = \underbrace{H_{\text{LoS}} + H_{\text{Env}}}_{H_S} + \underbrace{H_{\text{ref}_1} + \dots + H_{\text{ref}_n}}_{\text{body parts: } H_B}, \quad (1)$$

where H_{LoS} and H_{Env} are denoted as static path H_S , and H_{ref_i} is the reflection path from an arbitrary body part. For simplicity, all the reflection paths from body parts are denoted as H_B . The signal representation is shown in Fig. 2(c). Obviously, when the sleep posture slightly changes, the positions of the body parts in 3D space change accordingly. This change will cause the reflection paths change (i.e., H_B), which can be considered as another *unique feature* for 3D sleep posture recognition.

3.3 The Complementarity Between Blocked Tags and Unblocked Tags

To better understand how these two unique features are pertinent to 3D sleep postures, we conduct a benchmark experiment on several typical examples. For simplicity, we let a user lie in the bed and purposely change sleep postures. The deployment details and system implementation are described in Section 5.1. The experimental results are shown in Fig. 3 and our observations are as follows.

- All the blocked tags illustrate a “clear” imprint of the human body in the 2D plane. When the sleep posture changes, the blocked tags are changed according to the relative location of the tag, body, and antenna. As shown in Fig. 3(a)-(f), for each sleep posture, both the RSSI image and phase image depict the imprint of this posture in the 2D plane.
- The signals of those unblocked tags can provide the third dimension information of the body parts in the 3D space. When the user changes the position of her left leg in the 3D space, the multipath reflections from the leg will change accordingly. As shown in Fig. 3(g) and (h), we observe an obvious RSSI/phase variation

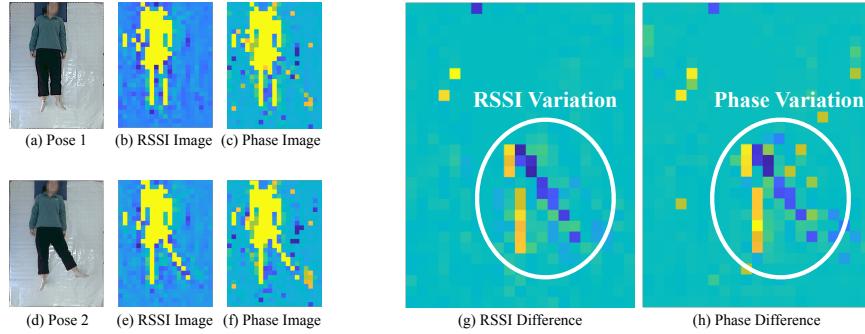


Fig. 3. (a)-(f) The sleep postures and their corresponding RSSI/phase images. (g)-(h) The RSSI/phase difference between these two sleep postures.

region that is around the left leg. This is because different body parts' positions in the 3D space lead to different multipath reflections. Hence, the signal variation of the nearby unblocked tags provides a good opportunity to acquire the third dimension information of the body part.

Therefore, the blocked tags and unblocked tags are complementary for sensing the sleep postures in the 3D space, which can be used as a reliable and unique primitive for 3D sleep posture recognition. Besides, to better represent the 2D body imprint, we set the value of blocked tags as 0, and the unblocked tags as 1, obtaining a binary matrix. Similarly, we denote this binary matrix as the *binary image*.

4 DESIGN

In this section, we first introduce the system overview of TagSleep3D, and then elaborate on the design details.

4.1 System Overview

TagSleep3D can recognize the 3D sleep posture skeleton using RF signals. It takes the signal features captured by the RFID reader as input and returns the 3D positions of 14 keypoints of the human body. As shown in Fig. 4, TagSleep3D consists of the following modules that operate in a pipelined manner to train a deep learning model for 3D sleep posture recognition.

- Data collection: when a person is sleeping in the bed, the RFID reader continuously collects the backscattered responses including the RSSI and phase from the tag array. Meanwhile, a camera captures the video streams simultaneously for obtaining the groundtruth labels of the 3D sleep posture. The setup details are introduced in Section 5.1.
- Data pre-processing: for the raw measured RSSI and phase readings, TagSleep3D first filters out the noise imposed by the hardware and the environment. Then, it removes the body motions and keeps the stable data segments to produce the RF images (Section 3.1). After that, TagSleep3D pre-processes the RF images including denoising, upsampling, and normalization (Section 4.2).
- Data augment: two data augmentation methods are proposed to enrich the dataset including the diversity of deployment, bed size, and sleep postures (section 4.3).
- 3D label generation: we propose a light-weight 3D pose label generation method using only one non-wearable camera. By recognizing the 2D pose first, TagSleep3D can provide reliable 3D pose labels for training the model. (Section 4.4)
- Network model: the 3D sleep posture recognition model takes the processed RF images and their binary images as input and employs the attention mechanism into the convolutional neural network. This model finally outputs the 3D positions of 14 keypoints of the human body (Section 4.5).

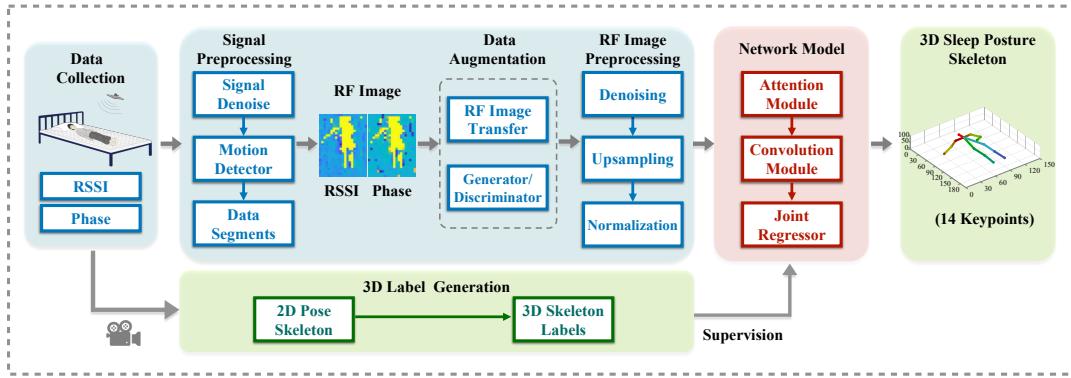


Fig. 4. Overview of TagSleep3D.

Once the model is trained, we can use this model to infer the 3D sleep posture skeletons from RF signals alone during online recognition. Note that for a stable sleep posture that lasts for a period, our model can predict a sequence of 3D pose skeletons to reduce the side effects caused by noises.

4.2 Data Pre-processing

The original RF images may contain noise and are not ready for 3D sleep posture recognition. Thus, we need to preprocess the RF images for TagSleep3D, including the RF image denoising, upsampling, and normalization.

4.2.1 Denoising. Since TagSleep3D employs a large number of tags (e.g., 588 tags for a queen bed and 728 tags for a king bed), tag collisions and interference may occur and thus some tags may not respond to the reader occasionally. Consequently, when forming RF images, some non-blocked tags may be mistakenly identified as blocked tags, resulting in noises in the RF images. These random noises can affect the accuracy of 3D sleep posture recognition. For simplicity, we denote these tags as noisy tags. As shown in Fig. 5, the noisy tags are these randomly scattered pixels (yellow) on the original RSSI image and phase image.

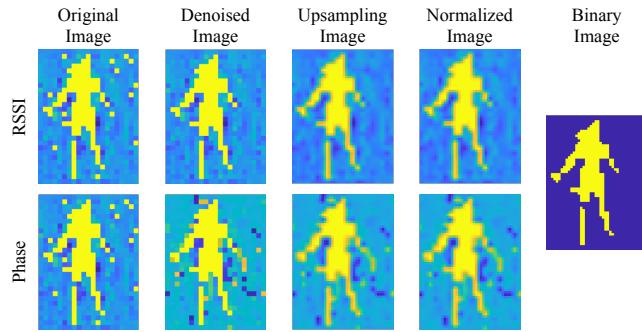


Fig. 5. Preprocessing RSSI image and phase image.

To remove the scattered pixels and avoid their side effects, we first identify the noisy tags and then give them proper RSSI and phase values, converting them from blocked tags to non-blocked tags. Specifically, we utilize the 8-connectivity method to identify all connected areas formed by blocked tags. We then select the largest connected area as the human body and its center as the body center. We calculate the shortest distance from each connected area to the body center. When the distance is smaller than a threshold (e.g., 95 cm according to the average torso length), the tags in that area should be normally blocked tags caused by body occlusion;

otherwise, the tags in that area are all noisy tags. Note that if the number of tags in a connected area is one, we directly identify this tag as the noisy tag. Eventually, for each noisy tag, its RSSI/phase value can be given by the average RSSI/phase values of its 8-neighborhood tags (i.e., up, down, left, right, top-left, top-right, bottom-left, and bottom-right). The original RSSI/phase images and the denoised are also shown in Fig. 5. With this method, most noisy tags can be assigned with a proper value, which is sufficient for our 3D sleep posture recognition. Note that as the number of tags in the array increases beyond our current maximum of 728, the number of noisy tags will increase significantly. In this case, we can employ multiple readers/antennas or modify the scheduling algorithm [7, 51] to mitigate the collision and reduce the number of noisy tags.

4.2.2 Upsampling. Each tag can be considered as a pixel in the RF image. Given a bed size, the number of pixels (e.g., 588) is relatively lower compared to a typical image captured by the camera. The low-resolution RF images make it more challenging in extracting fine-grained 3D sleep postures. To improve the resolution of the RF images, we employ a bilinear interpolation method to obtain 2× upsampled RF images.

4.2.3 Normalization. In order to eliminate the influence of scale differences of the input data and facilitate better model training and generalization, we normalize the RSSI and phase values separately. We apply the min-max method and the normalized RSSI and phase images can be obtained by

$$r_{m,n}^* = \frac{r_{m,n} - r_{\min}}{r_{\max} - r_{\min}}, \quad (2)$$

$$p_{m,n}^* = \frac{p_{m,n} - p_{\min}}{p_{\max} - p_{\min}}, \quad (3)$$

where, $r_{m,n}$ and $p_{m,n}$ are the RSSI and phase values of tag $(m, n)^{th}$, and r_{\min} (p_{\min}) and r_{\max} (p_{\max}) are the minimum and maximum RSSI (phase) values in the RSSI (phase) image, respectively. Fig. 5 shows the original RF images and the corresponding preprocessed images, respectively.

4.3 Data Augmentation

The performance of 3D sleep posture recognition can be affected by various factors, such as device deployment, bed size, and the diversity of sleep postures. To resolve this issue, we present two data augmentation methods to enrich the dataset and bolster data variability, ultimately enhancing the robustness and adaptability of TagSleep3D.

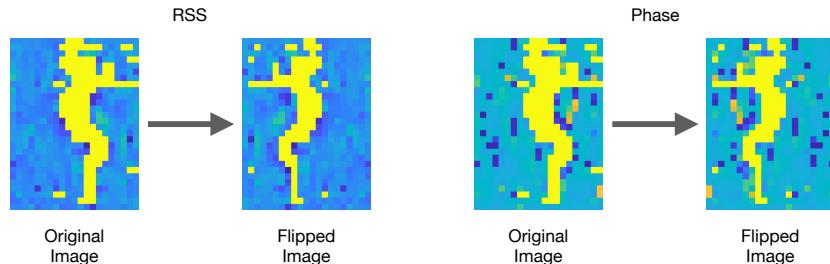


Fig. 6. The original RF images and the flipped RF images.

4.3.1 Augmenting the diversity of deployment and bed size. To enhance the diversity of bed size and antenna placement, we propose a data augmentation method based on RF image transformation, including the RF image flipping, cropping, and filling.

Flipping. In practical deployment, it is impossible to maintain the same position of the antenna in different bedrooms. For example, the antenna can be normally deployed on the left or right side of the bed depending on the specific bedroom layout. Hence, we consider the scenario where the antenna placement is mirrored to

the opposite side of the bed, while its relative positions to the tag array and the human body are unchanged. To maintain consistent distances between the antenna and body limbs after this mirroring, we also mirror the sleep posture. To emulate the above scenario, we can flip the original RF images horizontally. Fig. 6 illustrates the original RF images and the flipped RF images, respectively.

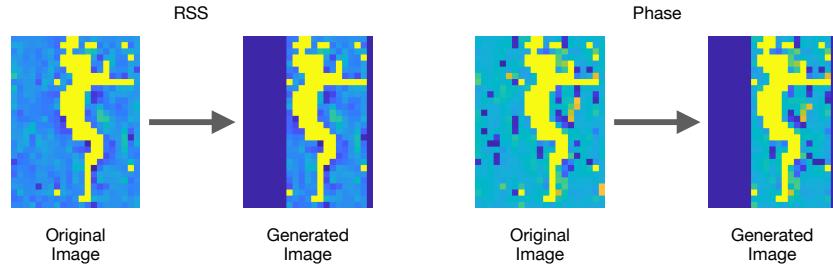


Fig. 7. The original RF images (queen size) and the generated RF images (other size).

Cropping and filling. When the bed size changes, the number of tags deployed on the mattress also changes. Hence, the dimension of RF images changes accordingly. To make TagSleep3D adaptable to different bed sizes, we resize the original collected RF images to generate new RF images for different bed sizes. In TagSleep3D, the RF images were collected on a typical queen-size bed and we denote these RF images as the original RF images. Specifically, we emulate the scenario where the bed size varies while other factors remain unchanged, such as the sleep posture and the spatial relationships between the antenna, the tag array, and the human body. To simulate smaller bed sizes, we need to crop the original RF images properly. Since the sleep posture and the spatial relationships are unchanged for the target bed size, we can directly crop the original RF images based on the identified boundaries of the human body, using the connected graph method, to fit the target bed size. Additionally, for all training data, the dimensions of the input RF images should be identical. We then standardize the cropped RF image size by filling in its removed area with a constant value (i.e., the minimum RSSI and phase values of the non-blocking tags) to enlarge it to the queen size. The original RF images and the generated images are shown in Fig. 7. When adapting our trained model for use in larger beds, we remove data outside the body boundaries to adjust the RF image to the standard queen size. This approach eliminates the need to create RF images for larger bed sizes during data augmentation.

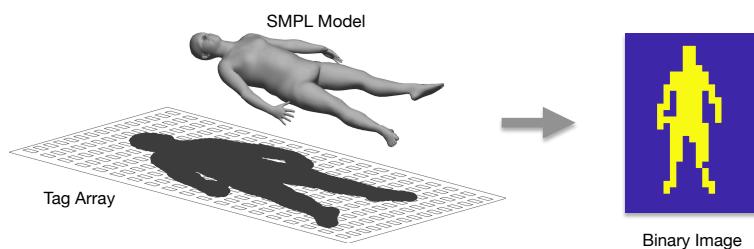


Fig. 8. Converting the SMPL sleep posture into the binary image.

4.3.2 Augmenting the diversity of sleep postures. Sleep postures are diverse and it is therefore time-consuming to collect a large number of training samples with rich sleep postures. Therefore, we employ the SLP sleep posture dataset [55] to enrich our dataset. Although SLP is a computer vision-based posture dataset, we only use their pose annotations which offer valuable information about sleep postures. The pose annotation is provided by the

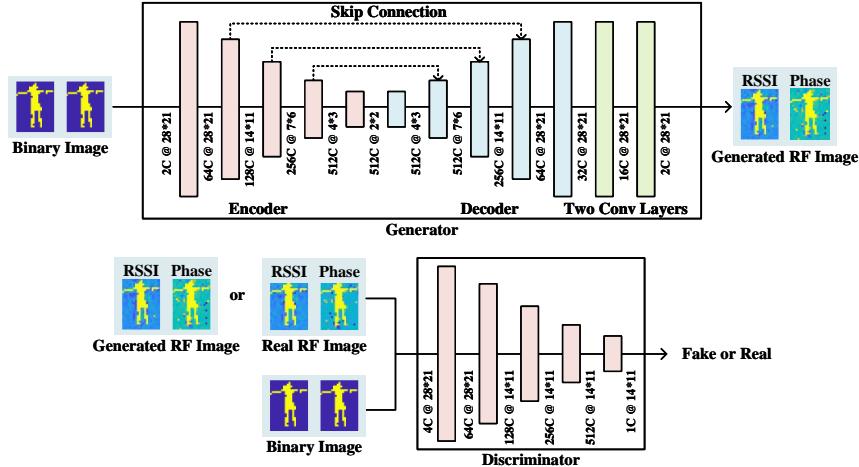


Fig. 9. The structure of the pix2pix network. The encoder has four downsampling layers, and each layer is composed of a 2D convolutional layer (Conv2d), a batch normalization layer (BatchNorm2d), and an activation function (LeakyReLU). In contrast, the decoder has four upsampling layers, and each layer is composed of a 2D transposed convolutional layer (ConvTranspose2d), a BatchNorm2d, and a LeakyReLU.

SMPL (Skinned Multi-Person Linear) model [15] which is a 3D representation of the human body containing 24 joints and various body parameters. Recall that we use the RF images (RSSI and phase images) and their binary images to infer the 3D sleep postures. Therefore, we need to establish a connection between the pose annotations in the SLP dataset and our RF images or binary images. Accordingly, our data augmentation method consists of two parts. The first part is to transfer the pose annotations (i.e., the SMPL model) from the SLP dataset into our binary images. For simplicity, we denote these binary images as SLP-binary images. The second part involves using these SLP-binary images to generate corresponding RF images, including RSSI and phase images.

For the first part, as shown in Fig. 8, we align the center of the SMPL model with the center of our RFID tag array. This alignment ensures that the binary image reflects how the human body would interact with the tags. We then generate an SLP-binary image by marking pixels as ‘1’ if the corresponding tag is completely obscured by the SMPL model and ‘0’ otherwise. This step simulates how different sleep postures would physically block the RFID tags. The second part is to transfer these SLP-binary images into the desired RF images. To do so, we employ a conditional generative adversarial network, specifically the pix2pix network, which takes a binary image as input and outputs the corresponding RF image. We train this network with our collected dataset, including binary images and their corresponding real RF images from various users and sleep postures. The pix2pix network consists of a generator and a discriminator. The generator aims to convert the input binary images into RF images, making the generated RF images as much similar to the real RF images as possible. The discriminator is to distinguish the differences between the generated RF images and the real RF images. Once the pix2pix network is trained, we can use the generator to generate the desired RSSI/phase images from any given SLP-binary image.

Pix2pix network design. The structures of the generator and discriminator are shown in Fig. 9. The generator consists of a U-net structure and two convolutional layers. The U-net structure divides the generator into an encoder and a decoder, with three skip connections between the encoder and decoder. The skip connections can preserve the sleep posture information in the binary image, making the generated RF images more realistic. After the encoder and decoder, two convolutional layers are used to improve the generator’s ability to learn the relationship between binary images and RF images. Note that, in the encoder, Conv2d is used to compress and

extract features from the input binary image; BatchNorm2d can accelerate the training speed; LeakyReLU can improve the stability of the training process. In the decoder, ConvTranspose2d is used to upsample features and convert low-dimensional features back into high-dimensional features. To ensure the same feature dimension, padding is applied to the output features at the end of each upsampling layer.

For the discriminator, the Pix2pix network uses PatchGAN which consists of five convolutional layers and each layer is composed of a Conv2d, a BatchNorm2d, and a LeakyReLU. The inputs of the discriminator are the real binary images paired with either the real RF image or the generated RF image. After five convolutional layers, PatchGAN maps the input data into a 14×11 matrix. For each element in the matrix, we calculate the probability that this element is real and we use the mean probability as the final output of the discriminator.

Recall that the optimization goal of the pix2pix network is to minimize the difference between the generated RF images and real RF images, while maximizing the accuracy in distinguishing real RF images and generated RF images. Hence, the objective function can be expressed as

$$\min_G \max_D \mathcal{L}(G, D) = \mathbb{E}_{x,y} [\log D(x, y)] + \mathbb{E}_x [1 - \log D(x, G(x))], \quad (4)$$

where x represents the binary image and y is the real RF image. $G(x)$ denotes the RF image generated by the input binary image. $D(x, y)$ represents the probability that the discriminator identifies the real RF image as true, and $D(x, G(x))$ corresponds to the probability that the discriminator considers the generated RF image as true.

To better minimize the difference between generated RF images and real RF images, we further use the mean absolute error (L1 loss) to optimize the generator. The L1 loss function is calculated as follows:

$$\mathcal{L}_{L1}(G) = \mathbb{E}_{x,y} [\|y - G(x)\|_1]. \quad (5)$$

Therefore, the final optimization goal of the pix2pix network can be expressed as:

$$G^* = \min_G \max_D \mathcal{L}(G, D) + \lambda \mathcal{L}_{L1}(G), \quad (6)$$

where λ is the coefficient of the L1 loss function. In the training process, the generator and the discriminator are trained alternately using Adam optimizer [43] based on our collected data.

4.4 3D Skeleton Label Generation

In TagSleep3D, 3D skeleton labels of sleep postures are needed for training the model with supervised learning. Existing methods employ wearable cameras [89] or multiple cameras [105] deployed at different positions to label 3D skeletons, which are not convenient in sleep scenarios. Instead, we use a single non-wearable Microsoft Kinect 2.0 [38] to obtain 3D labels. Note that once our model is trained, the Kinect is no longer needed for online recognition. Our system then independently infers the 3D sleep posture skeletons from RF signals. The Kinect is synchronized with the RFID device and captures both RGB images and depth images. Although the software development kit for Kinect can extract 3D human skeletons, the accuracy of this method is low in sleep scenarios. This is mainly because when a person lies on the bed, his/her body is closely connected to the background, making it difficult to distinguish the body pose from the background by depth information alone. To solve this problem, we propose to extract 2D skeleton keypoints from Kinect's RGB images first, and then combine the depth values to generate the 3D pose skeleton labels.

Specifically, we first remove the RGB images and depth images that contain body motions according to RF signals. Next, we leverage the OpenPose system [9] which takes the processed RGB images as input and outputs 14 keypoints of the 2D human skeleton. We also manually corrected those very few keypoints that OpenPose failed to detect. Lastly, for each 2D keypoint, we map it to the corresponding depth image to find its third-dimensional value, thereby generating the 3D sleep posture skeleton. Note that since the depth data may contain noise, we use the mean depth value calculated within a vicinity of 3×3 pixels for each keypoint.

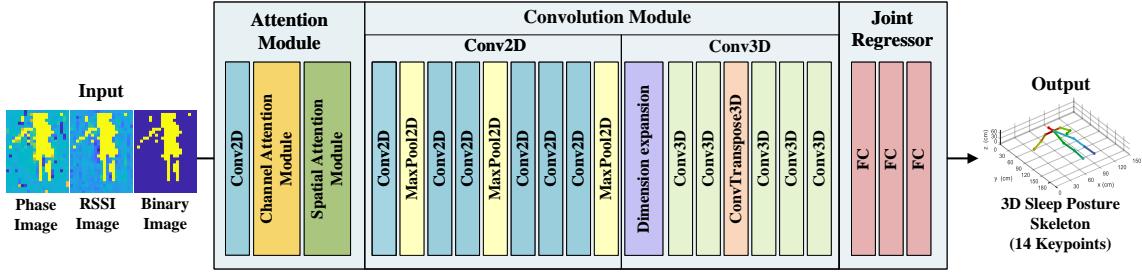


Fig. 10. The network structure. The attention module is composed of a channel attention module and a spatial attention module. The convolution module is composed of 2D convolution layers and 3D convolution layers. The joint regressor is composed of three fully connected layers.

4.5 Network Model

So far we have obtained many labeled sleep posture examples and their corresponding RSSI images, phase images, and binary images. Intuitively, we can manually design some features based on the RF images. However, such features should consider the reflection, the posture-related information, and the constraints on the locations of body joints, which is actually a challenging task. Alternatively, we select the convolutional neural network that is promising for complex tasks to automatically extract features together with the attention mechanism to leverage the complementarity between the blocked and unblocked tags. Thus, our goal is to design a 3D sleep posture recognition network that takes the RSSI images, phase images, and binary images as input data and outputs the 3D sleep posture skeleton. Note that the 3D skeleton labels obtained from Section 4.4 serve as the supervision data and the training dataset is augmented using the method described in Section 4.3.

When designing the network, we need to take into account the cross-user scenarios. This is because different users have different body shapes, which affects the wireless signal differently. Recall that we treat the signals from the tag array as RF images, which inherently capture the user's body shape. For instance, taller users tend to have longer body imprints in the RF images, while shorter users have smaller imprints. Consequently, the input RF images enable the network to output 3D skeletons that also reflect individuals' body shapes. Additionally, we collect diverse data and ensure that the users in our testing dataset are unseen during the training phase. If the initial performance is not satisfactory, we iteratively adjust the network structure or fine-tune its parameters. This diverse data enables the network to learn from a wide range of body shapes and sizes, enhancing its ability to generalize well across different users. The final network structure is described below.

The network model consists of three parts: the attention module, convolution module, and joint regressor (see Fig. 10). The attention module extracts useful features while suppressing the irrelevant information by multiplying the attention map and the input feature map. The convolution module further extracts features, and adds a new dimension to the 2D features. The joint regressor eventually maps the 3D features to the 3D sleep posture skeleton.

4.5.1 Attention module. We use a 2D convolution layer to output 2D features F from the input RSSI images, phase images, and binary images. Then, the 2D features F are input to the channel attention module to extract useful information and output features F' . Next, F' is used as the input of the spatial attention module to focus on more meaningful features [93].

4.5.2 Convolution module. The convolution module is used to optimize the features and convert the 2D features into 3D features. Naively, the 2D features can be directly extended to 3D by using repetition, but this method may bring in ambiguities. Instead, we employ the method in [59] to add a new channel to the extended 3D features, providing the height of each voxel. Specifically, for 2D features, we use a 2D convolutional neural network



Fig. 11. An example of TagSleep3D deployment.

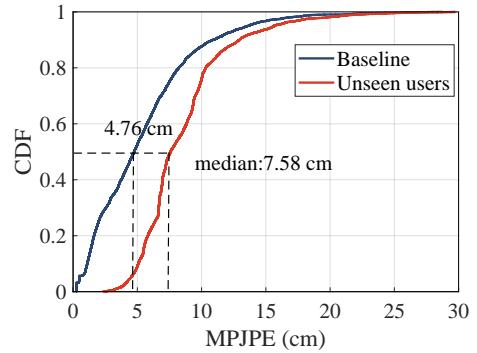


Fig. 12. MPJPE of baseline and unseen users.

consisting of six 2D convolutional layers with a kernel size of 3×3 , a stride of 1, and a padding of 0. In each 2D convolutional layer, we use both the activation function LeakyReLU and batch normalization layer BatchNorm2d to accelerate the training speed and improve its performance. Next, the 2D features are expanded to a new dimension, transforming them into 3D features. We then add a new channel to the 3D features. Finally, we use a 3D convolutional neural network to further optimize the 3D features, especially the sleep posture-related features. The 3D convolutional neural network includes four 3D convolutional layers, one 3D transposed convolution layer, and one 3D convolutional layer. Each layer has a kernel size of $3 \times 3 \times 3$, a stride of 1, and padding of 1.

4.5.3 Joint regressor. To predict the positions and connections of the body joints, we design a joint regressor that combines the extracted 3D features with the human skeleton structure. The joint regressor consists of three fully connected layers. Specifically, the fully connected layer is directly connected to the unfolded 3D features, extracting the joint information of the 3D sleep postures. The activation function ReLU is then used in the end for outputting the recognized 3D sleep posture skeletons.

4.5.4 Loss function. The joint 3D position loss function is widely used to calculate the difference between the predicted result and the true result. In TagSleep3D, the joint 3D position loss function is defined as the average Euclidean distance between the predicted joint position and the true joint position, which can be calculated as

$$L_{\text{joints}} = \frac{\sum_{i=1}^n \|x_i - \hat{x}_i\|_2^2}{n}, \quad (7)$$

where n is the number of joints, x_i the true joint position, and \hat{x}_i the predicted joint position. A smaller L_{joints} indicates higher accuracy of the predicted model. Thus, during the training process, we optimize our model by minimizing L_{joints} using the Adam optimizer [43].

5 EVALUATION

In this section, we evaluate the performance of TagSleep3D in terms of effectiveness and robustness through various experiments. We also compare TagSleep3D with state-of-the-art works.

5.1 Implementation

System setup. The hardware of TagSleep3D mainly includes an Impinj R420 RFID reader with a fixed frequency of 924.375 MHz, a circularly polarized antenna with a gain of 8 dBi, and a tag array. The antenna is connected to the reader through a UHF coaxial feeder and the reader is connected to a laptop via an Ethernet cable. The reader can obtain backscattered signals from the tags at a sampling rate of 3 to 5 reading samples per tag per

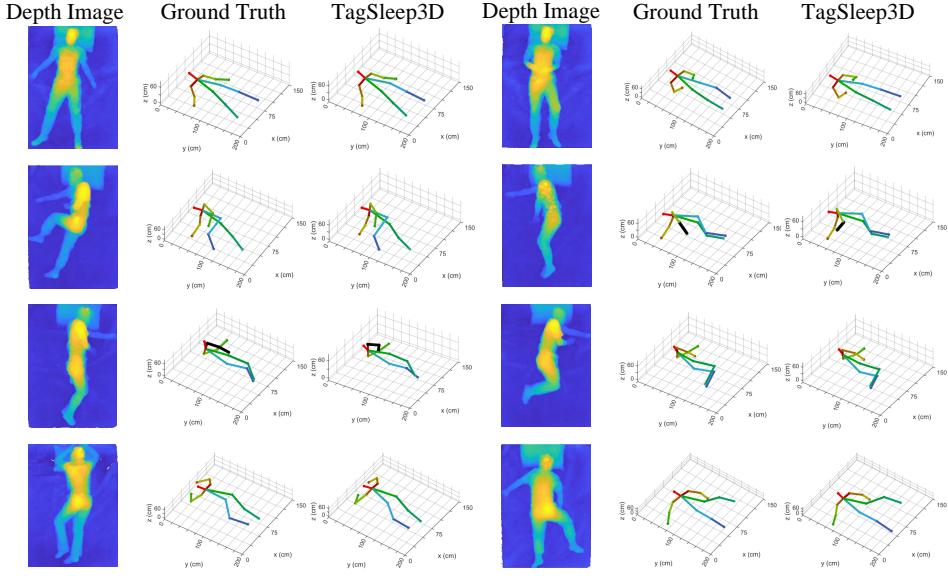


Fig. 13. Examples of the estimated 3D sleep posture results.

time window (i.e., 3 seconds). We deploy a Microsoft Kinect V2 [38] camera connected to another laptop to collect ground-truth depth/RGB images at a frame rate of 30 fps. We use the Network Time Protocol (NTP) to synchronize the RFID device and Kinect. Specifically, we timestamp all RSSI/phase measurements and Kinect frames and synchronize these two data streams according to their respective timestamps. The ground-truth 3D skeleton labels are obtained using the method proposed in Section 4.4.

The tag array consists of a number of tags to fully cover the bed. For convenience, all tags are first attached to a bed-size plastic film, and then we place the plastic film under the bed sheet to avoid the shifting caused by human movements during sleep. Note that given a specific type of tag, its size is fixed, and hence, deploying fewer tags would induce larger sensing gaps. To reduce the sensing gap and also avoid the coupling effect between tags, we set the interval between two adjacent tags as 2 cm according to [51]. Taking a twin-size bed and the H47 tags (44×44 mm) as an example, there are totally 28×21 tags in the tag array². We hang the antenna from the ceiling and it can be placed either on the left or the right side of the bed. The distance between the antenna and tag array is flexible, varying from 1.6 m to 2.3 m depending on the specific bedroom layout. The Kinect camera is also deployed right above the bed and there is also no strict constraint for its position. One example of the experimental setup is shown in Fig. 11.

Data collection and participants. We evaluated TagSleep3D and collected sleep posture data with 43 users, including 31 males and 12 females, aged from 19 and 60 years old, weighing between 40 and 92 Kg, and height from 155 to 185 cm. During the experiment, the users sleep in bed with their natural sleep postures. For instance, they can turn over and change their sleep postures during sleep. We have totally collected 27,300 sleep posture samples varying from different bedrooms, different bed sizes, and different hardness levels of mattresses.

Metrics. We employ the following widely used metrics to evaluate the system performance: (i) L2 distance, which is the Euclidean distance between the estimated joint and the ground truth, (ii) Mean Per Joint Position Error (MPJPE), and (iii) Average joint localization error.

²Note that using a smaller tag can further increase the resolution of the tag array, but the reading range would be decreased with smaller tags, especially for high-density deployment.

Table 1. Average joint localization errors (cm).

Joint	Head	Neck	RSho	RElb	RWri	LSho	LElb	LWri	RHip	RKne	RAnk	LHip	LKne	LA nk	Aver
X	3.59	2.62	1.79	2.39	3.41	1.99	4.42	2.44	2.29	4.95	4.62	3.52	6.61	7.46	3.72
Y	4.12	4.06	3.04	2.21	6.08	5.76	4.29	5.52	9.52	4.57	3.62	3.09	2.33	3.75	4.43
Z	3.74	1.99	4.04	4.76	1.85	3.65	3.06	3.99	3.57	3.28	2.71	5.45	4.59	1.66	3.45

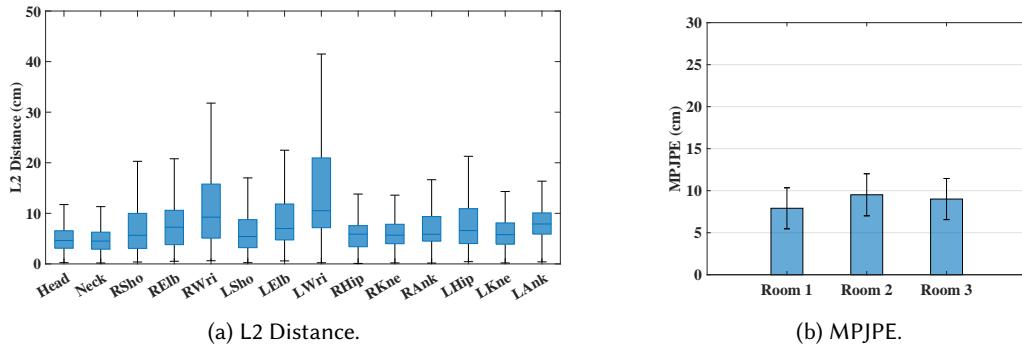


Fig. 14. Overall performance.

5.2 Overall Performance

For the overall performance, we first evaluate TagSleep3D in cross-user scenarios. Then, we evaluated TagSleep3D in terms of the network model, the data augmentation, and the 3D label generation.

5.2.1 Cross-user scenarios. For the baseline, we randomly divide each user’s data into two parts with a proportion of 8:2. The 80% of the data from all 43 users are used for model training and the remaining 20% is the testing dataset. For cross-user scenarios, we trained the model with data from randomly selected 30 users and tested it using data from 13 users not included in the training phase. Note that the training dataset also includes the generated dataset from our data augmentation methods. Fig. 12 illustrates the CDF of the estimated error. For the baseline, TagSleep3D achieves a median MPJPE of 4.76 cm, indicating high accuracy in estimating the 3D skeleton. For unseen users, the median MPJPE is 7.58 cm. While there is a slight increase compared to the baseline, it still demonstrates satisfactory accuracy, which reveals that our model can generalize well to unseen users. In the following evaluations, we consistently make sure that the users in the testing dataset are not included in the training dataset, i.e., evaluating the system performance with unseen users.

5.2.2 Evaluation of 3D sleep posture recognition. Here, we show the overall performance of TagSleep3D on the 3D sleep posture recognition for unseen users. Fig. 13 illustrates several examples of the estimated sleep posture skeletons as well as the ground truth postures and depth images extracted from the Kinect camera. As we can see, our estimated 3D sleep posture skeletons are consistent with the ground-truth skeletons and the depth images for various sleep postures. To further evaluate the accuracy of each joint, we calculated the average joint localization error. As shown in Table 1, the average error is 3.72 cm, 4.43 cm, and 3.45 cm for the x, y, and z axes, respectively, which indicates the high performance of TagSleep3D. Furthermore, Fig. 14a demonstrates the L2 distance of each joint. As we can see, the average error varies slightly among different joints. For instance, the L2 distance of the head and neck (i.e., median error of 4.6 cm and 4.5 cm) is relatively lower compared to the wrists and ankles (i.e., median error of 9.2 cm and 7.8 cm). This is mainly because the wrists and ankles are small body joints against the body, and they are often co-located with other body parts in many sleep postures, e.g., hands on stomach or thighs.

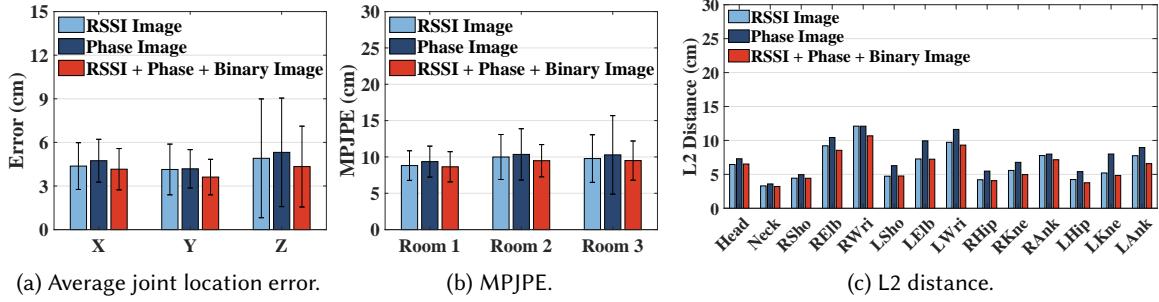


Fig. 15. The performance under different model inputs.

In addition, Fig. 14b shows the MPJPE in different bedrooms. Clearly, TagSleep3D can achieve satisfying performance in all three bedrooms with an average MPJPE of 7.9 cm, 9.5 cm, and 9.0 cm, respectively. It is worth noting that bedroom 2 and bedroom 3 exhibit a slightly higher error compared to bedroom 1. The reason is that the mattresses used in these two rooms are softer and smaller, which may affect the RF images. We further provide a detailed discussion in Section 5.3.

5.2.3 Evaluation of model inputs. An ablation study has been performed upon different inputs of our network model, including inputs such as RSSI images only, phase images only, and full input of TagSleep3D, i.e., a combination of RSSI images, phase images, and binary images. We first present the results of the average joint error on the x, y, and z axes, as shown in Fig. 15a. Compared with TagSleep3D (4.16 cm, 3.62 cm, 4.34 cm), we observe an average error increases of 10.67% and 17.26% when presenting the model with RSSI images only (4.37 cm, 4.13 cm, 4.91 cm) and phase images only (4.74 cm, 4.18 cm, 5.31 cm), respectively. Next, we evaluate the model input when switching bedrooms as shown in Fig. 15b. Similarly, we can see an average error increase of 3.3% and 8.08% with RSSI images only (i.e., 8.82 cm, 9.99 cm, and 9.77 cm for 3 bedrooms) and phase images only (9.35 cm, 10.35 cm, and 10.28 cm), compared to TagSleep3D (8.64 cm, 9.48 cm, and 9.50 cm). Finally, in Fig. 15c, we present the evaluation of L2 distance on the 14 joints, and on average, we notice error increases of 5.7% and 26.19% for RSSI Images only and phase images only, respectively. Overall, the input of RSSI Images only performs better than the phase Images only, given the robustness of RSSI to environmental noise. Although noisy, phase images provide rich information on postures (e.g., body reflections). As such, the combination of RSSI images, phase images, and binary images, i.e., TagSleep3D, yields the least error across different settings.

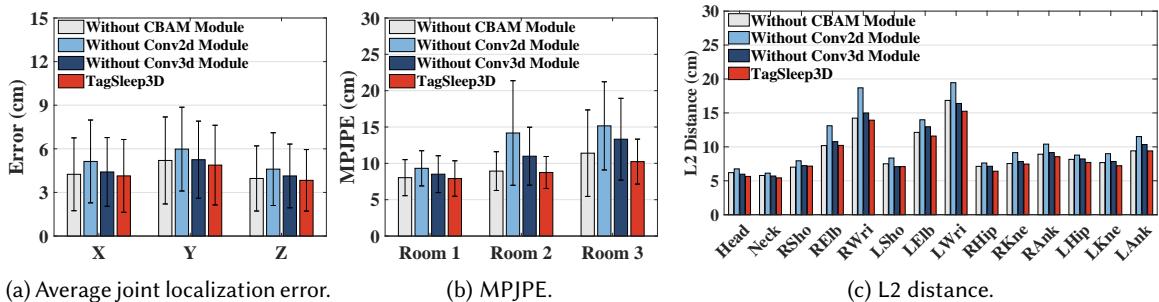


Fig. 16. The performance under different model structures.

5.2.4 Evaluation of model structures. We conducted an ablation study on the modules of the 3D sleep posture recognition model, namely, the attention mechanism module, the 3D convolution module, and the 2D convolution module. Fig. 16 presents the results of removing these modules one at a time, with TagSleep3D representing the complete model incorporating all modules. Furthermore, to mitigate the influence of bed size and hardness, we

included different bed sizes and hardness levels in this study. In Fig. 16a, we observe that the average joint error on the x, y, and z axes increases by 3.99%, 21.34%, and 7.05% after removing the attention mechanism module, the 3D convolution module, and the 2D convolution module, respectively. Similarly, we have average error increases of MPJPE across different bedrooms by 4.62%, 40.5%, and 19.88%, as shown in Fig. 16b, and of L2 distance by 4.37% 20.25%, and 6.32%, as shown in Fig. 16c. As indicated by the result, the 2D convolution module is more crucial in the model, considering its ability in exacting essential spatial features of images.

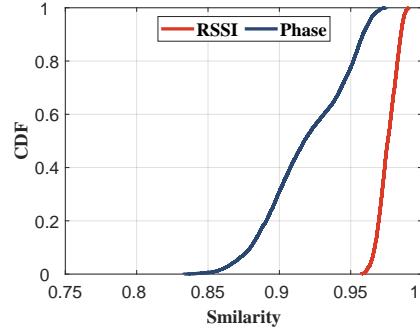


Fig. 17. The similarity between the original RF images and the generated RF images.

5.2.5 Evaluation of data augmentation. In order to validate the quality of the generated data, we calculate the similarity between the real RF images and the generated RF images from the pix2pix network. Here, we adopt the Cosine similarity which is a common metric for evaluating the similarity between two matrices [75]. Typically, the value is closer to 1, the two matrices are more similar to each other. As illustrated in Fig. 17, the similarity of 50% of examples is larger than 0.98 for RSSI and 0.94 for phase, respectively. The results indicate that the generated RF images exhibit a high degree of similarity to the real RF images, making them suitable for training and improving the performance of the 3D sleep posture recognition network.

To further evaluate the effectiveness of data augmentation, we employ four different datasets as the training dataset, respectively. The first dataset is our collected data; the second dataset is the collected data plus the generated data using RF image flipping (i.e., denoted as Gen. 1), which can enhance the diversity of the antenna deployment; the third dataset is the second one plus the generated data using RF image cropping and padding (i.e., Gen. 1 + Gen. 2), enriching the diversity of bed sizes; the last dataset is the third one plus the generated data using the pix2pix network (i.e., Gen. 1 + Gen. 2 + Gen. 3) which augments the diversity of the sleep postures.

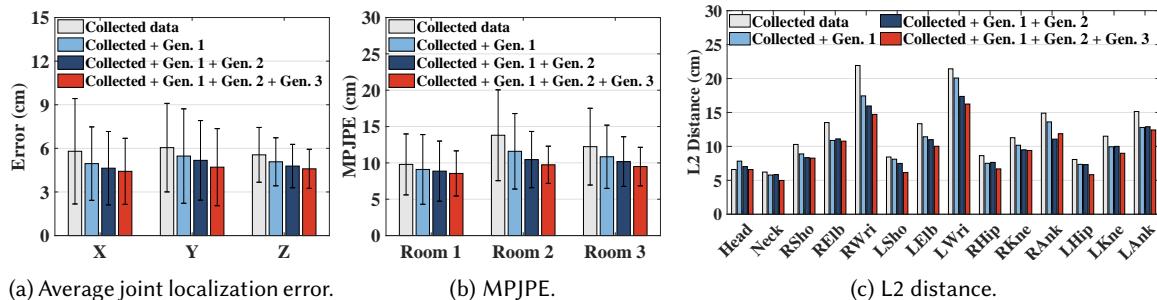
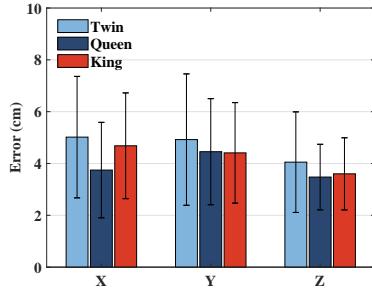


Fig. 18. The performance under different training datasets.

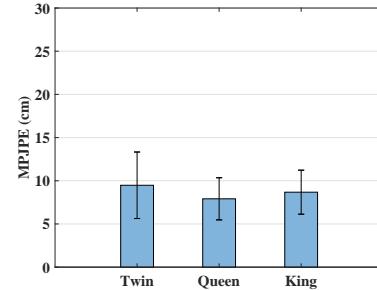
The results of the ablation study are depicted in Fig. 18. Specifically, Fig. 18a illustrates the average joint error on the x, y, and z axes when using different datasets, resulting in average error increases of 21.1%, 10.17%, and 4.96% compared to TagSleep3D which was trained with the fourth dataset. Additionally, Fig. 18b presents MPJPE

Table 2. The average joint localization errors of the generated 3D labels (cm).

Joint	Head	Neck	RSho	RElb	RWri	LSho	LElb	LWri	RHip	RKne	RArk	LHip	LKne	LAk	Aver
X	1.9	2.2	2.9	2.1	2.4	3.3	2.3	2.3	3.3	1.7	2.1	3.8	1.8	2.0	2.44
Y	1.9	2.7	2.3	2.2	2.2	2.6	1.9	2.0	2.7	2.9	3.0	2.7	2.6	2.7	2.46
Z	1.3	1.1	1.2	1.4	1.7	1.2	1.3	2.2	1.5	1.1	1.3	1.6	1.0	1.3	1.37



(a) Average joint location error along X, Y, and Z axes.



(b) MPJPE.

Fig. 19. Impact of bed size.

across different bedrooms, indicating increases of 21.5%, 10.02%, and 4.68%, respectively. Lastly, Fig. 18c represents the L2 distance, leading to increases of 21.16%, 11.45%, and 6.79%, respectively. According to the result, using the collected data as the foundation, the augmentation of our data can enrich its diversity and effectively enhance the robustness and accuracy of our model in real-world scenarios.

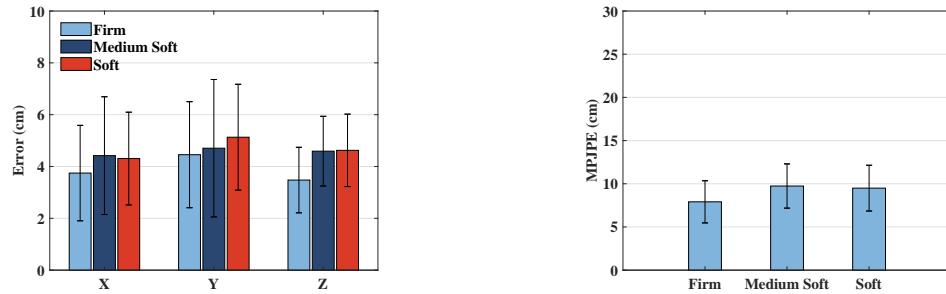
5.2.6 Evaluation of 3D skeleton labels. To evaluate the accuracy of 3D skeleton labels generated by our method, we randomly select 350 sleep posture visual images and manually annotate them. Compared to these manually labeled ground truths, the mean absolute error of our method is $4.35 \text{ cm} \pm 1.3 \text{ cm}$. Table 2 shows the average joint localization errors of different joints along the x, y, and z axes, respectively. The average errors of our 3D skeleton labels are 2.44 cm, 2.46 cm on the horizontal plane, and 1.37 cm along the vertical axis. These results demonstrate that our 3D labels are reasonably accurate and sufficient for TagSleep3D to serve as the ground truth in supervised training.

5.3 Impact of Various Factors

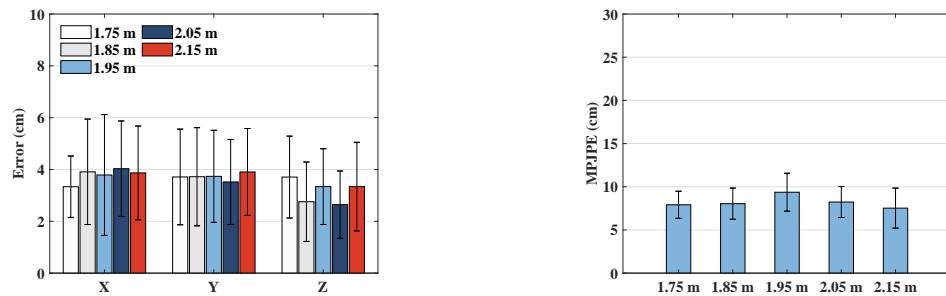
Next, we evaluate the robustness of TagSleep3D under various factors, such as the bed size, hardness level of the mattress, antenna height, and blanket.

5.3.1 Impact of bed size. Recall that the size of the tag array varies as the bed size changes, resulting in different dimensions of RF images. To keep an identical dimension for the model input, we resize the collected RF images to fit the queen size using the method in Section 4.3.1. Fig. 19a illustrates the average joint location errors for three widely used bed sizes. As we can see, the average error is no more than 5 cm along the x, y, and z axes for all three bed sizes. We also observe that the average error for twin-size bed is slightly larger than the other two sizes (i.e., 0.77 cm and 0.43 cm, respectively). This is because when fewer tags are deployed on the bed, the total amount of information caused by body reflection may decrease as well. In addition, as shown in Fig. 19b, the MPJPE are 9.48 cm, 7.91 cm, and 8.68 cm for the twin, queen, and King-size beds, respectively, which is stable across different bed sizes, indicating that TagSleep3D is robust to bed size changes.

5.3.2 Impact of hardness level of mattress. For a fair comparison, we first fix the bed size as the queen size and then employ different hardness levels of the mattress to evaluate TagSleep3D. Fig. 20a and Fig. 20b demonstrate

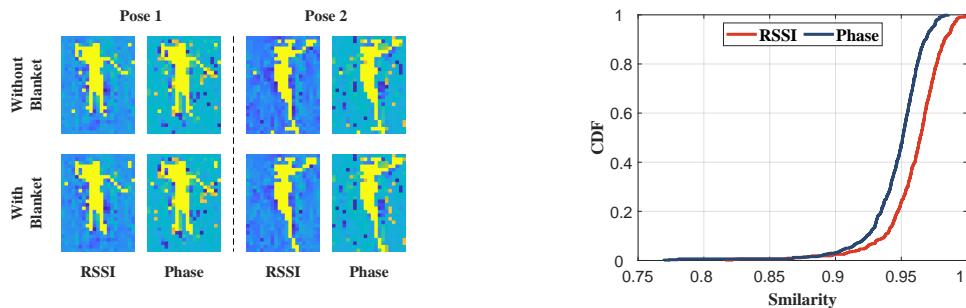


(a) Average joint location error along X, Y, and Z axes. (b) MPJPE.
Fig. 20. Impact of hardness level of mattress.



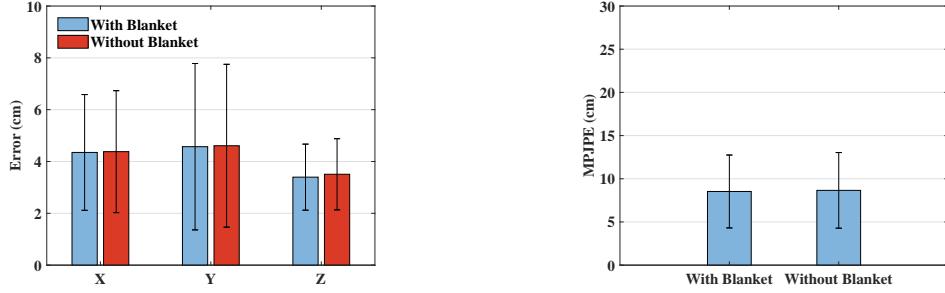
(a) Average joint location error along X, Y, and Z axes. (b) MPJPE.
Fig. 21. Impact of antenna height.

the average joint location error and MPJPE under three commonly used mattresses, i.e., the firm, medium soft, and soft. As we can see, the minimum (maximum) joint error and MPJPE are 3.5 cm (5.2 cm) and 7.9 cm (9.8 cm) for all three mattresses, which indicates that TagSleep3D is not sensitive to the hardness of the mattress though the tag array is deployed on the mattress. It is worth noting that both the average error and MPJPE slightly increase when the mattress is getting soft. The reason is that when a user sleeps on a soft mattress, his/her weight may cause some deformations on the mattress, which in turn affects the position and shape of tags.



(a) RSSI and phase images w/o blanket. (b) The similarity w/o blanket.
Fig. 22. The RF signals are similar w/o blanket.

5.3.3 Impact of antenna height. In TagSleep3D, the antenna height is flexible according to the bedroom layout. Given the communication range of RFID, as well as the typical bedroom layouts, we vary the antenna height

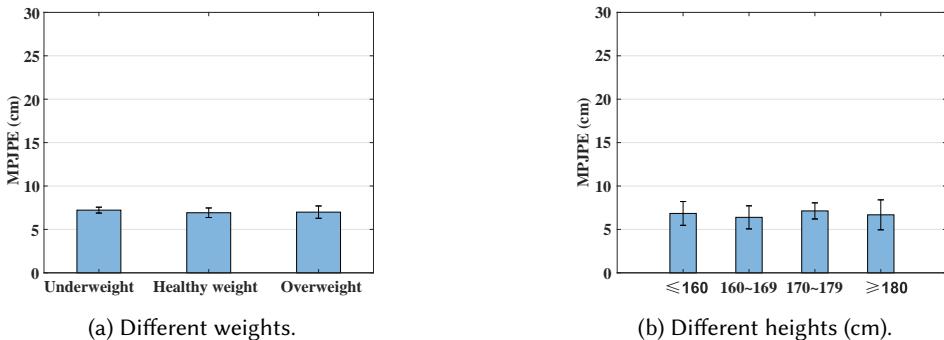


(a) Average joint location error along X, Y, and Z axes.
 Fig. 23. Impact of blanket.

(b) MPJPE.

from 1.75 m to 2.15 m with a step of 0.1 m. The average joint error and MPJPE are shown in Fig. 21a and Fig. 21b, respectively. Obviously, the performance of TagSleep3D is stable when the antenna height increases.

5.3.4 Impact of blankets. Before evaluating how blankets affect recognition accuracy, let's first take a look at the RF images when the user is covered by a blanket. Fig. 22a shows two examples of RSSI/phase images with and without the blanket, respectively. As we can see, both the RSSI and phase images with the blanket are similar to that without the blanket. To further quantify the similarity, we calculate the Cosine similarity again. Fig. 22b illustrates the cumulative distribution function (CDF) of the Cosine similarity with 320 pairs of RF images. For a fair comparison, each pair of examples was collected from the same person under the same sleep posture, while different pairs were from different persons under different sleep postures. Clearly, the similarity of 50% of examples is larger than 0.97 and 0.95 for RSSI and phase, respectively, indicating a strong similarity. Thus, the blanket does not have an obvious influence on the RF signals. To further verify the impact of blanket on the recognition accuracy, the testing samples w/o the blanket were collected with the same user under the same sleep posture for a fair comparison. As shown in Fig. 23a, the average joint error with the blanket is very similar to that without the blanket. Moreover, as can be seen from Fig. 23b, there is also no obvious difference in MPJPE between blanket and no blanket, i.e., the MPJPE of 8.53 cm and 8.66 cm, respectively. These results indicate that TagSleep3D can accurately recognize the 3D sleep postures when the user is covered by a blanket.



(a) Different weights.

(b) Different heights (cm).

Fig. 24. Impact of body weights and heights.

5.3.5 Impact of body weights and heights. For body weights, we divide the users into three categories based on their BMI (Body Mass Index): underweight ($BMI < 18.5$), healthy weight ($18.5 \leq BMI < 25$), and overweight ($BMI \geq 25$) [69]. To make a fair comparison, we select the users who slept in the same bedroom and on the same type of mattress. As shown in Fig. 24a, the overall error is stable across different weights with the MPJPE of 7.22 cm, 6.92 cm, and 6.99 cm for the underweight, healthy weight, and overweight, respectively. Then, we select the users

with healthy weight and divide their body heights into four categories according to [22]. As shown in Fig. 24b, the MPJPE of these four categories are 6.84 cm, 6.39 cm, 7.13 cm, and 6.68 cm, respectively, with a slightly larger error for the 170-179 cm category, which is mainly caused by the variations in sleep postures within this group. Nevertheless, the MPJPE is relatively stable across different body heights.

5.4 In Comparison to State-of-the-art Works

We compare TagSleep3D with the state-of-the-art 3D sleep posture recognition systems, including the visual systems, the pressure-mattress based systems, and the RF-based systems.

5.4.1 Compared to pressure and visual systems. Since the devices used in these systems are different, we summarize their reported results in Table 3. Note that most visual systems use the metric Percentage of Correct Keypoints (PCK), especially PCKh@0.5. Thus, we recalculate our results using PCKh@0.5. Clearly, TagSleep3D outperforms the pressure-mattress-based method with an error drop of 3.2 cm. Besides, it can achieve comparable accuracy with the multi-modal-based method. TagSleep3D can also achieve the PCKh@0.5 of 92.47%, which is slightly lower than the visual systems using RGB, long wave infrared (LWIR), and depth images, but our result is still satisfactory. This is reasonable since CV-based systems can obtain a higher-resolution image of the sleep posture.

Table 3. Comparison with the state-of-the-art systems.

System	Modality	MPJPE (cm)	PCKh@0.5
Bodies at Rest [14]	Pressure	11.18	-
Yu Yin et al. [100]	Multi-modal	7.06	-
Liu et al. [55]	RGB/LWIR/Depth	-	97.8%/96%/97.9%
MC-VAE [8]	LWIR+RGB	-	94.3%
TagSleep3D	RFID	7.58	92.47%

5.4.2 Compared to RF-based systems. Argosleep [2] is a 3D sleep posture monitoring system using mmWave devices. For seen users, Argosleep reported a median MPJPE of 6.22 cm without the height classifier, while our system outperforms Argosleep with a median MPJPE of 4.76 cm. However, when Argosleep employs the height classifier, it can improve the median MPJPE to 2.3 cm. When both seen and unseen users are included in the testing dataset, Argosleep reported a median MPJPE of 7.5 cm with the height classifier. Our system, which does not require height information, shows a slightly better performance with a median MPJPE of 6.58 cm³.

Next, we compare our system in 2D scenarios with TagSheet [51] which uses an RFID tag array to identify six common sleep postures: the supine, prone, left log, right log, left foetus, and right foetus. For a fair comparison, we filter out our dataset to align with these six posture types identified by TagSheet. Our system outputs a 3D sleep posture skeleton with 14 body joints. For classification, we analyze the differences in positions between the right and left shoulders (and hips) along the z-axis. If these differences exceed a threshold (e.g., 15 cm based on our dataset), we classify the posture as left log or left foetus; if negative and beyond the threshold, it's classified as right log or right foetus. For supine or prone postures, we compare the shoulder or hip positions along the x-axis. Additionally, we calculate the bending angles of the legs to distinguish between log and foetus postures. We then implement the hierarchical recognition approach proposed in TagSheet with a slight adjustment of some thresholds based on our collected sleep posture dataset.

The confusion matrix is shown in Fig. 25. As we can see, the accuracy of each sleep posture is above 98% for our system, while only in some cases, it misidentifies the supine as left log or prone. This is mainly caused by occasional estimation errors in the positions of shoulders or hips. For TagSheet, the recognition accuracy of

³For a fair comparison, we maintain the same ratio of seen and unseen users as Argosleep.

the left log, right log, and right foetus is higher than 92%, while the accuracy of other sleep postures is slightly lower. In some cases, it struggles to distinguish supine from prone postures and sometimes confuses them with lateral postures. This is due to its reliance on the distance between the body and head axes and the number of blocked tags (i.e., the number of pixel ‘1’ in the binary image), which can be misleading in certain postures, such as crossed hands or legs and heads facing to the side in supine postures. Note that the accuracy of TagSheet in our study is lower than that reported in their original research. This is mainly because our collected dataset includes more diverse and complex postures than the six standard postures that TagSheet was initially designed to identify. The above comparison results demonstrate that both TagSheet and our system can identify different types of sleep posture. Compared to TagSheet, our system recognizes fine-grained 3D skeletons before classifying the postures, and hence it can more reliably classify a wide range of sleep postures into their respective categories.

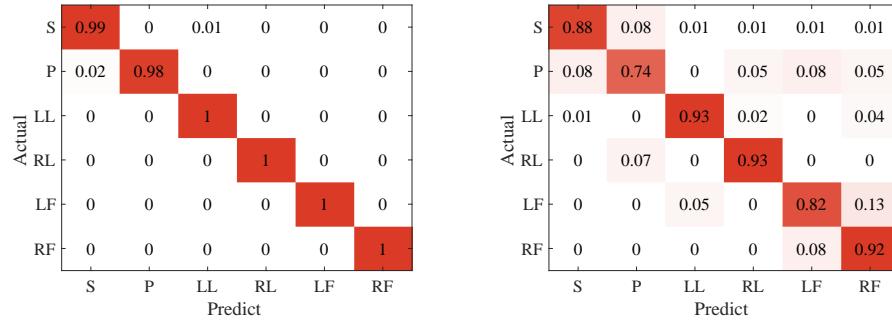


Fig. 25. The confusion matrix of sleep posture recognition. S, P, LL, RL, LF, and RF represent the supine, prone, left log, right log, left foetus, and right foetus, respectively.

6 DISCUSSIONS

We discuss some limitations and future works to further improve TagSleep3D as below.

Small body parts. Currently, the recognition accuracy of small body parts (e.g., the wrists and ankles) is not as high as other body joints. We can further employ cGan to generate more challenging samples like that or modify the weights for small body parts when training the model. We leave it for our future work.

Two users. TagSleep3D can be extended to recognize the 3D sleep postures of multiple users simultaneously. When two users, such as a couple, sleep in bed, we can first divide the RF images into two parts corresponding to their body positions. We then input the RF images of each user into TagSleep3D for recognition.

Human perspiration. It is common to see that people sweat during sleep. In fact, the RFID tags are covered by plastic film, making it resistant to human perspiration or mild humidity.

RF radiation. In TagSleep3D, we employ the RFID reader with a radiated power of 0.25 w and the antenna with 8 dBi gains. The distance between the antenna and the human body is around 2 m, which meets the FCC requirement and is safe for the user [4].

Non-static periods. TagSleep3D is adept at capturing coarse-grained movements such as turning over, showing acceptable performance in non-static periods with the MPJPE of 14.75 cm. To further refine our system, we can increase the sampling rate and enhance the quality of RF images. These improvements could enable our system to identify fine-grained motion procedures and improve overall accuracy. We leave it for our future work.

Data augmentation. TagSleep3D focus on the scenario where the antenna is moved from one side of the bed to the other side, which is a common setup in practice. For more challenging scenarios, such as moving the antenna to an arbitrary location, we can employ the diffusion model and transformers to generate more data with variations while ensuring high fidelity. We leave it for future work.

7 CONCLUSION

In this paper, we present a novel 3D sleep posture recognition system using RFID tags. With careful analyses of the relationship between the tag array and sleep postures, we propose a deep learning model composed of the attention mechanism and convolutional neural network to retrieve the 3D sleep postures by analyzing the RSSI/phase images from the tag array. We further propose two data augmentation methods to enrich the diversity of both sleep postures and environmental settings. Extensive experiments demonstrate that TagSleep3D can achieve a robust and high accuracy across users and environments.

ACKNOWLEDGMENTS

This work is supported by NSFC A3 Foresight Program under Grant 62061146001, National Natural Science Foundation of China (62272388, 62372372), Key Research and Development Program of Shaanxi (2024GH-YBXM-09), and the Shaanxi Qinchuangyuan Program under Grant QCYRCXM-2023-103.

REFERENCES

- [1] Aakriti Adhikari, Hem Regmi, Sanjib Sur, and Srihari Nelakuditi. 2022. MiShape: Accurate Human Silhouettes and Body Joints from Commodity Millimeter-Wave Devices. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 6, 3 (2022), 1–31.
- [2] Aakriti Adhikari and Sanjib Sur. 2023. Argosleep: Monitoring Sleep Posture from Commodity Millimeter-Wave Devices. In *IEEE INFOCOM 2023-IEEE Conference on Computer Communications*.
- [3] Dhruv Agarwal and Donna Thomas. 2020. Posture And Activity Analysis For Patients In Rehabilitation. *IJARIT* 10 (2020).
- [4] Darmendra D Arumugam and Daniel W Engels. 2008. Impacts of RF radiation on the human body in a passive RFID environment. In *2008 IEEE Antennas and Propagation Society International Symposium*. IEEE, 1–4.
- [5] Paolo Barsocchi. 2012. Position recognition to support bedsores prevention. *IEEE journal of biomedical and health informatics* 17, 1 (2012), 53–59.
- [6] Alexander Bigalke, Lasse Hansen, Jasper Diesel, Carlotta Hennigs, Philipp Rostalski, and Mattias P Heinrich. 2022. Anatomy-guided domain adaptation for 3D in-bed human pose estimation. *arXiv preprint arXiv:2211.12193* (2022).
- [7] Carlos Bocanegra, Mohammad Khojastepour, Mustafa Arslan, Eugene Chai, Sampath Rangarajan, and Kaushik Chowdhury. 2020. RFGo: a seamless self-checkout system for apparel stores using RFID. In *MobiCom 2020 - Proceedings of the Annual International Conference on Mobile Computing and Networking*. 1–14.
- [8] Ting Cao, Mohammad Ali Armin, Simon Denman, Lars Petersson, and David Ahmedt-Aristizabal. 2022. In-bed human pose estimation from unseen and privacy-preserving image domains. In *2022 IEEE 19th International Symposium on Biomedical Imaging (ISBI)*. IEEE, 1–5.
- [9] Zhe Cao, Gines Hidalgo, Tomas Simon, Shih-En Wei, and Yaser Sheikh. 2019. OpenPose: realtime multi-person 2D pose estimation using Part Affinity Fields. *IEEE transactions on pattern analysis and machine intelligence* 43, 1 (2019), 172–186.
- [10] Leslie Casas, Nassir Navab, and Stefanie Demirci. 2019. Patient 3D body pose estimation from pressure imaging. *International journal of computer assisted radiology and surgery* 14, 3 (2019), 517–524.
- [11] Liqiong Chang, Jiaqi Lu, Ju Wang, Xiaojiang Chen, Dingyi Fang, Zhanyong Tang, Petteri Nurmi, and Zheng Wang. 2018. SleepGuard: capturing rich sleep information using smartwatch sensing data. *Ubicomp* 2, 3 (2018), 98.
- [12] James Allan Cheyne. 2002. Situational factors affecting sleep paralysis and associated hallucinations: position and timing effects. *Journal of Sleep Research* 11, 2 (2002), 169–177.
- [13] Sudhansu Chokroverty et al. 2010. Overview of sleep & sleep disorders. *Indian J Med Res* 131, 2 (2010), 126–140.
- [14] Henry M Clever, Zackory Erickson, Ariel Kapusta, Greg Turk, Karen Liu, and Charles C Kemp. 2020. Bodies at rest: 3d human pose and shape estimation from a pressure image using synthetic data. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 6215–6224.
- [15] Henry M. M. Clever, Patrick Grady, Greg Turk, and Charles C. Kemp. 2022. BodyPressure - Inferring Body Pose and Contact Pressure from a Depth Image. *IEEE Transactions on Pattern Analysis and Machine Intelligence* (2022), 1–1. <https://doi.org/10.1109/TPAMI.2022.3158902>
- [16] Naima Covassin, Ariel Neikrug, Lianqi Liu, Jody Corey-Bloom, Jose Loredo, Barton Palmer, Jeanne Maglione, and Sonia Ancoli-Israel. 2012. Clinical correlates of periodic limb movements in sleep in Parkinson’s disease. *Journal of the neurological sciences* 316 (01 2012), 131–6.
- [17] Bappaditya Debnath, Mary O’Brien, Motonori Yamaguchi, and Ardhendu Behera. 2021. A review of computer vision-based approaches for physical rehabilitation and assessment. *Multimedia Systems* 28 (06 2021).

- [18] Fei Deng, Jianwu Dong, Xiangyu Wang, Ying Fang, Yu Liu, Zhaofei Yu, Jing Liu, and Feng Chen. 2018. Design and implementation of a noncontact sleep monitoring system using infrared cameras and motion sensor. *IEEE Transactions on Instrumentation and Measurement* 67, 7 (2018), 1555–1563.
- [19] Gustavo Desouzart, Ernesto Filgueiras, and Rui Matos. 2015. Relationship between postural reeducation technique during sleep and relaxation technique in sleep quality. *Procedia Manufacturing* 3 (2015), 6093–6100.
- [20] Han Ding, Zhenbin Chen, Cui Zhao, Fei Wang, Ge Wang, Wei Xi, and Jizhong Zhao. 2023. MI-Mesh: 3D Human Mesh Construction by Fusing Image and Millimeter Wave. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 7, 1 (2023), 1–24.
- [21] Fitbit. 2019. *Fitbit*. Retrieved May 10, 2019 from <https://www.fitbit.com>
- [22] Cheryl D Fryar, Margaret D Carroll, Qiuping Gu, Joseph Afful, and Cynthia L Ogden. 2021. Anthropometric reference data for children and adults: United States, 2015–2018. *U.S. Department of Health and Human Services, Centers for Disease Control and Prevention, National Center for Health Statistics* (January 2021).
- [23] Hulya Gokalp and Malcolm Clarke. 2013. Monitoring activities of daily living of the elderly and the potential for its use in telecare and telehealth: a review. *TELEMEDICINE and e-HEALTH* 19, 12 (2013), 910–923.
- [24] Claudia Gorecki, S José Closs, Jane Nixon, and Michelle Briggs. 2011. Patient-reported pressure ulcer pain: a mixed-methods systematic review. *Journal of pain and symptom management* 42, 3 (2011), 443–459.
- [25] Timo Grimm, Manuel Martinez, Andreas Benz, and Rainer Stiefelhagen. 2016. Sleep position classification from a depth camera using bed aligned maps. In *2016 23rd International Conference on Pattern Recognition (ICPR)*. IEEE, 319–324.
- [26] Reinder Haakma and R Beun. 2012. Unobtrusive sleep monitoring. In *Measuring Behavior*, Vol. 2012. 122.
- [27] Mehrdad Heydarzadeh, Mehrdad Nourani, and Sarah Ostadabbas. 2016. In-bed posture classification using deep autoencoders. In *2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. IEEE, 3839–3842.
- [28] Hillrom. 2023. Advancing connected care. Online. <https://www.hillrom.com>
- [29] Rong-Shue Hsiao, Zhenqiang Mi, Bo-Ru Yang, Lih-Jen Kau, Mekuanint Agegnehu Bitew, and Tzu-Yu Li. 2016. Body posture recognition and turning recording system for the care of bed bound patients. *Technology and Health Care* 24, s1 (2016), S307–S312.
- [30] Weimin Huang, Aung Aung Phyoe Wai, Siang Fook Foo, Jit Biswas, Chi-Chun Hsia, and Koujuchi Liou. 2010. Multimodal sleeping posture classification. In *2010 20th International conference on pattern recognition*. IEEE, 4336–4339.
- [31] Shang-Rung Hwang, Sheng-Wei Hwang, Jin-Cherng Chen, and Juen-Haur Hwang. 2019. Association of periodic limb movements during sleep and Parkinson disease: A retrospective clinical study. *Medicine* 98 (12 2019), e18444.
- [32] Prasara Jakkaew and Takao Onoye. 2020. Non-contact respiration monitoring and body movements detection for sleep using thermal imaging. *Sensors* 20, 21 (2020), 6307.
- [33] Jawbone. 2018. *Jawbone*. Retrieved Jul. 31, 2018 from <https://jawbone.com/>
- [34] Po-Yuan Jeng, Li-Chun Wang, Chaur-Jong Hu, and Dean Wu. 2021. A Wrist Sensor Sleep Posture Monitoring System: An Automatic Labeling Approach. *Sensors* 21, 1 (2021), 258.
- [35] Sanghoon Jeon, Anand Paul, Sang Hyuk Son, and Yongsoon Eun. 2018. Sleep position management system for enhancing sleep quality using wearable devices. In *Proceedings of the 16th ACM Conference on Embedded Networked Sensor Systems*. 365–366.
- [36] Wenjun Jiang, Hongfei Xue, Chenglin Miao, Shiyang Wang, Sen Lin, Chong Tian, Srinivasan Murali, Haochen Hu, Zhi Sun, and Lu Su. 2020. Towards 3D human pose construction using WiFi. In *Proceedings of the 26th Annual International Conference on Mobile Computing and Networking*. 1–14.
- [37] David A Johnson, William C Orr, Joseph A Crawley, Barry Traxler, Joseph McCullough, Kurt A Brown, and Thomas Roth. 2005. Effect of esomeprazole on nighttime heartburn and sleep quality in patients with GERD: a randomized, placebo-controlled trial. *American Journal of Gastroenterology* 100, 9 (2005), 1914–1922.
- [38] mobajemu-msft Karl-Bridge-Microsoft, alexbuckgit. 2023. *Kinect for Windows*. Retrieved August 31, 2022 from <https://learn.microsoft.com/en-gb/windows/apps/design/devices/kinect-for-windows>
- [39] Hiroki Kato, Yu Enokibori, Naoto Yoshida, and Kenji Mase. 2021. Toward Fine-Grained Sleeping Activity Recognition: 3d Extension and an Estimation Try on Joint Position of SLP Dataset. In *Adjunct Proceedings of the 2021 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2021 ACM International Symposium on Wearable Computers*. 322–327.
- [40] Adam Khan, Khoi D Than, Kevin S Chen, Anthony C Wang, Frank La Marca, and Paul Park. 2014. Sleep apnea and cervical spine pathology. *European Spine Journal* 23, 3 (2014), 641–647.
- [41] Ali Kiaghadi, Seyedeh Zohreh Homayounfar, Jeremy Gummesson, Trisha Andrew, and Deepak Ganeshan. 2019. Phyjama: physiological sensing via fiber-enhanced pyjamas. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 3, 3 (2019), 1–29.
- [42] Ali Kiaghadi, S Zohreh Homayounfar, Jeremy Gummesson, Trisha L Andrew, and Deepak Ganeshan. 2020. Phyjama: Physiological Sensing via Fiber-Enhanced Pyjamas. *GetMobile: Mobile Computing and Communications* 24, 1 (2020), 33–37.
- [43] Diederik P Kingma and Jimmy Ba. 2014. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980* (2014).

- [44] John E. Kiriazi, Shekh M. M. Islam, Olga Borić-Lubecke, and Victor M. Lubecke. 2021. Sleep Posture Recognition With a Dual-Frequency Cardiopulmonary Doppler Radar. *IEEE Access* 9 (2021), 36181–36194. <https://doi.org/10.1109/ACCESS.2021.3062385>
- [45] Anastasi Kosmadopoulos, Charli Sargent, David Darwent, Xuan Zhou, and Gregory D Roach. 2014. Alternatives to polysomnography (PSG): a validation of wrist actigraphy and a partial-PSG system. *Behavior research methods* 46, 4 (2014), 1032–1041.
- [46] Chung-Hsien Kuo, Fang-Chung Yang, Ming-Yuan Tsai, and Ming-Yih Lee. 2004. Artificial neural networks based sleep motion recognition using night vision cameras. *Biomedical Engineering: Applications, Basis and Communications* 16, 02 (2004), 79–86.
- [47] Jaehoon Lee, Min Hong, and Sungyong Ryu. 2015. Sleep monitoring system using kinect sensor. *International Journal of Distributed Sensor Networks* 11, 10 (2015), 875371.
- [48] Yan-Ying Li, Yan-Jing Lei, Lyn Chao-Ling Chen, and Yi-Ping Hung. 2018. Sleep posture classification with multi-stream CNN using vertical distance map. In *2018 International Workshop on Advanced Image Technology (IWAIT)*. IEEE, 1–4.
- [49] Wen-Hung Liao and Chien-Ming Yang. 2008. Video-based activity and movement pattern analysis in overnight sleep studies. In *2008 19th International Conference on Pattern Recognition*. IEEE, 1–4.
- [50] Chen Liu, Jie Xiong, Lin Cai, Lin Feng, Xiaojiang Chen, and Dingyi Fang. 2019. Beyond respiration: Contactless sleep sound-activity recognition using RF signals. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 3, 3 (2019), 1–22.
- [51] Jia Liu, Xingyu Chen, Shigang Chen, Xiulong Liu, Yanyan Wang, and Lijun Chen. 2019. TagSheet: Sleeping posture recognition with an unobtrusive passive tag matrix. In *IEEE INFOCOM 2019-IEEE Conference on Computer Communications*. IEEE, 874–882.
- [52] Jian Liu, Yingying Chen, Yan Wang, Xu Chen, Jerry Cheng, and Jie Yang. 2018. Monitoring vital signs and postures during sleep using WiFi signals. *IEEE Internet of Things Journal* 5, 3 (2018), 2071–2084.
- [53] Jason J Liu, Wenyao Xu, Ming-Chun Huang, Nabil Alshurafa, Majid Sarrafzadeh, Nitin Raut, and Behrooz Yadegar. 2013. A dense pressure sensitive bedsheet design for unobtrusive sleep posture monitoring. In *2013 IEEE international conference on pervasive computing and communications (PerCom)*. IEEE, 207–215.
- [54] Jason Jun Hing Liu. 2014. *Body Pressure Image Analyses: Static, Dynamic, and Sensing Strategies for a Pressure Sensitive Bedsheet*. Ph.D. Dissertation. UCLA.
- [55] Shuangjun Liu, Xiaofei Huang, Nihang Fu, Cheng Li, Zhongnan Su, and Sarah Ostadabbas. 2023. Simultaneously-Collected Multimodal Lying Pose Dataset: Enabling In-Bed Human Pose Monitoring. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 45, 1 (2023), 1106–1118. <https://doi.org/10.1109/TPAMI.2022.3155712>
- [56] Xuefeng Liu, Jiannong Cao, Shaojie Tang, and Jiaqi Wen. 2014. Wi-sleep: Contactless sleep monitoring via wifi signals. In *2014 IEEE Real-Time Systems Symposium*. IEEE, 346–355.
- [57] Xuefeng Liu, Jiannong Cao, Shaojie Tang, Jiaqi Wen, and Peng Guo. 2015. Contactless respiration monitoring via off-the-shelf WiFi devices. *IEEE Transactions on Mobile Computing* 15, 10 (2015), 2466–2479.
- [58] Chengwen Luo, Zhongru Yang, Xingyu Feng, Jin Zhang, Hong Jia, Jianqiang Li, Jiawei Wu, and Wen Hu. 2021. RFaceID: Towards RFID-based Facial Recognition. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 5, 4 (2021), 1–21.
- [59] Yiyue Luo, Yunzhu Li, Michael Foshey, Wan Shou, Pratyusha Sharma, Tomás Palacios, Antonio Torralba, and Wojciech Matusik. 2021. Intelligent Carpet: Inferring 3D Human Pose from Tactile Signals. In *2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. 11250–11260. <https://doi.org/10.1109/CVPR46437.2021.01110>
- [60] Manuel Martinez, Boris Schauerte, and Rainer Stiefelhagen. 2013. “BAM!” Depth-based Body Analysis in Critical Care. In *International Conference on Computer Analysis of Images and Patterns*. Springer, 465–472.
- [61] Akshay Menon and Manoj Kumar. 2013. Influence of body position on severity of obstructive sleep apnea: a systematic review. *International Scholarly Research Notices* 2013 (2013).
- [62] Vangelis Metsis, Dimitrios Kosmopoulos, Vassilis Athitsos, and Fillia Makedon. 2014. Non-invasive analysis of sleep patterns via multimodal sensor input. *Personal and ubiquitous computing* 18, 1 (2014), 19–26.
- [63] Edwin A Mitchell, Bradley T Thach, John MD Thompson, and Sheila Williams. 1999. Changing infants’ sleep position increases risk of sudden infant death syndrome. *Archives of pediatrics & adolescent medicine* 153, 11 (1999), 1136–1141.
- [64] Sara Mahvash Mohammadi, Majdi Alnowami, Sofia Khan, Derk-Jan Dijk, Adrian Hilton, and Kevin Wells. 2018. Sleep posture classification using a convolutional neural network. In *2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. IEEE, 1–4.
- [65] Zena Moore, Seamus Cowman, and Ronán M Conroy. 2011. A randomised controlled clinical trial of repositioning, using the 30 tilt, for the prevention of pressure ulcers. *Journal of clinical nursing* 20, 17-18 (2011), 2633–2644.
- [66] Jingyi Ning, Lei Xie, Chuyu Wang, Yanling Bu, Fengyuan Xu, Da-Wei Zhou, Sanglu Lu, and Baoliu Ye. 2023. RF-Badge: Vital Sign-Based Authentication via RFID Tag Array on Badges. *IEEE Transactions on Mobile Computing* 22, 2 (2023), 1170–1184.
- [67] Sleep Number. 2023. Adjustable and smart beds, bedding and pillows. Online. <https://www.sleepnumber.com>
- [68] Cecilia Occhiuzzi and Gaetano Marrocco. 2009. The RFID technology for neurosciences: feasibility of limbs’ monitoring in sleep diseases. *IEEE Transactions on Information Technology in Biomedicine* 14, 1 (2009), 37–43.
- [69] World Health Organization. 2023. *Body Mass Index (BMI)*. <https://www.who.int/data/gho/data/themes/topics/topic-details/GHO/body-mass-index> Accessed: October 14, 2023.

- [70] Sarah Ostadabbas, Maziyar Baran Pouyan, Mehrdad Nourani, and Nasser Kehtarnavaz. 2014. In-bed posture classification and limb identification. In *2014 IEEE Biomedical Circuits and Systems Conference (BioCAS) Proceedings*. IEEE, 133–136.
- [71] Aifeng Ren, Binbin Dong, Xiangyu Lv, Tianqiao Zhu, Fangming Hu, and Xiaodong Yang. 2016. A non-contact sleep posture sensing strategy considering three dimensional human body models. In *2016 2nd IEEE International Conference on Computer and Communications (ICCC)*. IEEE, 414–417.
- [72] Yili Ren, Zi Wang, Sheng Tan, Yingying Chen, and Jie Yang. 2021. Winect: 3D human pose tracking for free-form activity using commodity WiFi. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 5, 4 (2021), 1–29.
- [73] Yili Ren, Zi Wang, Yichao Wang, Sheng Tan, Yingying Chen, and Jie Yang. 2022. Gopose: 3d human pose estimation using wifi. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 6, 2 (2022), 1–25.
- [74] ResT. 2023. ResT smart bed technology – the world's first smart mattress. Online. <https://restperformance.com>
- [75] Amit Singhal et al. 2001. Modern information retrieval: A brief overview. *IEEE Data Eng. Bull.* 24, 4 (2001), 35–43.
- [76] Edward M Sitar IV and Sanjib Sur. 2022. A Millimeter-Wave Wireless Sensing Approach for Sleep Posture Classification. In *Proceedings of the 20th ACM Conference on Embedded Networked Sensor Systems*. 794–796.
- [77] Eight Sleep. 2023. The intelligent sleep system. Online. <https://www.eightsleep.com>
- [78] Soo-Jung So, Heon-Jeong Lee, Seung-Gul Kang, Chul-Hyun Cho, Ho-Kyoung Yoon, and Leen Kim. 2015. A comparison of personality characteristics and psychiatric symptomatology between upper airway resistance syndrome and obstructive sleep apnea syndrome. *Psychiatry investigation* 12, 2 (2015), 183.
- [79] Harun SÜMBÜL and Ahmet Hayrettin YÜZER. 2016. 3D Monitoring of lying position for patients with positional sleep apnea syndrome. *Journal of New Results in Science* 5 (2016), 59–70.
- [80] Wei Sun. 2021. RFitness: Enabling Smart Yoga Mat for Fitness Posture Detection with Commodity Passive RFIDs. In *2021 IEEE International Conference on RFID (RFID)*. 1–8.
- [81] Xiao Sun, Li Qiu, Yibo Wu, Yeming Tang, and Guohong Cao. 2017. Sleepmonitor: Monitoring respiratory rate and body position during sleep using smartwatch. *Ubicomp* 1, 3 (2017), 104.
- [82] Keison Tang, Arjun Kumar, Muhammad Nadeem, and Issam Maaz. 2021. CNN-Based Smart Sleep Posture Recognition System. *IoT* 2, 1 (2021), 119–139. <https://doi.org/10.3390/iot2010007>
- [83] Boonsin Tangtrakulwanich and Anucha Kapkird. 2012. Analyses of possible risk factors for subacromial impingement syndrome. *World journal of orthopedics* 3, 1 (2012), 5.
- [84] Tomed. 2021. *Somnobelt*. Retrieved 2021 from <https://en.tomed.com/products/somnobelt/>
- [85] Carlos Torres, Victor Fragozo, Scott D Hammond, Jeffrey C Fried, and BS Manjunath. 2016. Eye-cu: Sleep pose classification for healthcare using multimodal multiview data. In *2016 IEEE Winter conference on applications of computer vision (WACV)*. IEEE, 1–9.
- [86] Kenji Uchino, Makoto Shiraishi, Keita Tanaka, Masashi Akamatsu, and Yasuhiro Hasegawa. 2017. Impact of inability to turn in bed assessed by a wearable three-axis accelerometer on patients with Parkinson's disease. *PloS one* 12, 11 (2017), e0187616.
- [87] MA Van Herwaarden, David A Katzka, AJPM Smout, M Samsom, M Gideon, and DO Castell. 2000. Effect of different recumbent positions on postprandial gastroesophageal reflux in normal subjects. *The American journal of gastroenterology* 95, 10 (2000), 2731–2736.
- [88] Ellen R Van Kesteren, J Peter van Maanen, Anthony AJ Hilgevoord, D Martin Laman, and Nico de Vries. 2011. Quantitative effects of trunk and head position on the apnea hypopnea index in obstructive sleep apnea. *Sleep* 34, 8 (2011), 1075–1081.
- [89] Vicon. 2021. *Motion capture systems*. Retrieved 2021 from <https://www.vicon.com>
- [90] Chuyu Wang, Jian Liu, Yingying Chen, Lei Xie, Hong Bo Liu, and Sanclu Lu. 2018. RF-kinect: A wearable RFID-based approach towards 3D body movement tracking. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 2, 1 (2018), 1–28.
- [91] Chuyu Wang, Lei Xie, Wei Wang, Yingying Chen, Yanling Bu, and Sanglu Lu. 2018. RF-ECG: Heart Rate Variability Assessment Based on COTS RFID Tag Array. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 2 (07 2018), 1–26.
- [92] Shuai Wang, Dongjiang Cao, Ruofeng Liu, Wenchao Jiang, Tianshun Yao, and Chris Xiaoxuan Lu. 2023. Human Parsing with Joint Learning for Dynamic mmWave Radar Point Cloud. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 7, 1 (2023), 1–22.
- [93] Sanghyun Woo, Jongchan Park, Joon-Young Lee, and In So Kweon. 2018. CBAM: Convolutional Block Attention Module. In *Proceedings of the European Conference on Computer Vision (ECCV)*.
- [94] Weiye Xu, Jianwei Liu, Shimin Zhang, Yuanqing Zheng, Feng Lin, Jinsong Han, Fu Xiao, and Kui Ren. 2021. RFace: Anti-Spoofing Facial Authentication Using COTS RFID. In *IEEE INFOCOM 2021 - IEEE Conference on Computer Communications*. 1–10.
- [95] Xiaowei Xu, Feng Lin, Aosen Wang, Yu Hu, Ming-Chun Huang, and Wenya Xu. 2016. Body-earth mover's distance: A matching-based approach for sleep posture recognition. *IEEE transactions on biomedical circuits and systems* 10, 5 (2016), 1023–1035.
- [96] Xiaowei Xu, Feng Lin, Aosen Wang, Chen Song, Yu Hu, and Wenya Xu. 2015. On-bed sleep posture recognition based on body-earth mover's distance. In *2015 IEEE Biomedical Circuits and Systems Conference (BioCAS)*. IEEE, 1–4.
- [97] Fan Yang, Ren Li, Georgios Georgakis, Srikrishna Karanam, Terrence Chen, Haibin Ling, and Ziyang Wu. 2020. Robust Multi-modal 3D Patient Body Modeling. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, 86–95.

- [98] Lei Yang, Qiongzheng Lin, Xiangyang Li, Tianci Liu, and Yunhao Liu. 2015. See through walls with COTS RFID system!. In *MobiCom 2015 - Proceedings of the 21st Annual International Conference on Mobile Computing and Networking*. 487–499.
- [99] Zhicheng Yang, Parth H Pathak, Yunze Zeng, Xixi Liran, and Prasant Mohapatra. 2017. Vital sign and sleep monitoring using millimeter wave. *ACM Transactions on Sensor Networks (TOSN)* 13, 2 (2017), 1–32.
- [100] Yu Yin, Joseph P. Robinson, and Yun Fu. 2022. Multimodal In-Bed Pose and Shape Estimation under the Blankets. In *Proceedings of the 30th ACM International Conference on Multimedia* (Lisboa, Portugal) (MM '22). Association for Computing Machinery, New York, NY, USA, 2411–2419. <https://doi.org/10.1145/3503161.3548063>
- [101] Meng-Chieh Yu, Huan Wu, Jia-Ling Liou, Ming-Sui Lee, and Yi-Ping Hung. 2012. Multiparameter sleep monitoring using a depth camera. In *International Joint Conference on Biomedical Engineering Systems and Technologies*. Springer, 311–325.
- [102] Shichao Yue, Yuzhe Yang, Hao Wang, Hariharan Rahul, and Dina Katabi. 2020. Bodycompass: Monitoring sleep posture with wireless signals. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 4, 2 (2020), 1–25.
- [103] Zhiqiang Zhang and Guang-Zhong Yang. 2015. Monitoring cardio-respiratory and posture movements during sleep: What can be achieved by a single motion sensor. In *2015 IEEE 12th International Conference on Wearable and Implantable Body Sensor Networks (BSN)*. IEEE, 1–6.
- [104] Mingmin Zhao, Tianhong Li, Mohammad Abu Alsheikh, Yonglong Tian, Hang Zhao, Antonio Torralba, and Dina Katabi. 2018. Through-wall human pose estimation using radio signals. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 7356–7365.
- [105] Mingmin Zhao, Yonglong Tian, Hang Zhao, Mohammad Abu Alsheikh, Tianhong Li, Rumen Hristov, Zachary Kabelac, Dina Katabi, and Antonio Torralba. 2018. RF-based 3D skeletons. In *Proceedings of the 2018 Conference of the ACM Special Interest Group on Data Communication*. 267–281.
- [106] Zhihao Zhou, Sean Padgett, Zhixiang Cai, Giorgio Conta, Yufen Wu, Qiang He, Songlin Zhang, Chenchen Sun, Jun Liu, Endong Fan, et al. 2020. Single-layered ultra-soft washable smart textiles for all-around ballistocardiograph, respiration, and posture monitoring during sleep. *Biosensors and Bioelectronics* 155 (2020), 112064.
- [107] Kaiyin Zhu, T Douglas Bradley, Maryam Patel, and Hisham Alshaer. 2017. Influence of head position on obstructive sleep apnea severity. *Sleep and Breathing* 21, 4 (2017), 821–828.