Article

Exposing Data Leakage in Wi-Fi CSI-Based Human Action Recognition: A Critical Analysis

Domonkos Varga (ID)

Nokia Bell Labs, 1082 Budapest, Hungary; domonkos.varga@nokia-bell-labs.com

Abstract: Wi-Fi channel state information (CSI)-based human action recognition systems have garnered significant interest for their non-intrusive monitoring capabilities. However, the integrity of these systems can be compromised by data leakage, particularly when improper dataset partitioning strategies are employed. This paper investigates the presence and impact of data leakage in three published Wi-Fi CSI-based human action recognition methods that utilize deep learning techniques. The original studies achieve precision rates of 95% or higher, attributed to the lack of human-based dataset splitting. By re-evaluating these systems with proper subject-based partitioning, our analysis reveals a substantial decline in performance, underscoring the prevalence of data leakage. This study highlights the critical need for rigorous dataset management and evaluation protocols to ensure the development of robust and reliable human action recognition systems. Our findings advocate for standardized practices in dataset partitioning to mitigate data leakage and enhance the generalizability of Wi-Fi CSI-based models.

Keywords: data leakage; Wi-Fi CSI; human action recognition; deep learning; machine learning integrity

1. Introduction

In the ever-evolving landscape of technology, the intersection of wireless communication and human behavior analysis has given rise to an innovative field known as Wi-Fi signal-based human action recognition. This paradigm exploits the ubiquitous Wi-Fi signals that permeate our surroundings to decipher and interpret human actions, providing a non-intrusive and privacy-preserving approach to understanding and responding to human behavior [1]. Wi-Fi signal-based human action recognition involves the utilization of wireless communication signals, specifically Wi-Fi signals, to infer and understand human actions in a given environment. Traditional methods of human action recognition often rely on cameras and sensors [2], which may pose privacy concerns. Wi-Fi signals, on the other hand, are omnipresent and can penetrate walls, allowing for non-intrusive and remote monitoring of human actions [3]. This technology leverages the fluctuations in the Wi-Fi signal's channel state information (CSI) caused by human movements [4]. CSI refers to the information about the wireless channel that the Wi-Fi signal travels through, including phase, amplitude, and frequency characteristics [5]. By analyzing these variations, algorithms can deduce intricate details about human actions such as walking, running, sitting, or even gestures [6].

Wi-Fi signal-based human action recognition may find applications in the following areas:

1. Healthcare monitoring: Wi-Fi signal-based systems can be deployed in healthcare settings to monitor the movements of patients, especially the elderly, providing valuable insights into their daily activities and well-being [7].
2. Smart homes: Wi-Fi signal-based human action recognition can enhance automation systems by recognizing specific gestures to control devices, adjusting lighting, or regulating temperature based on occupants' activities [8].

3. Security and surveillance: Wi-Fi signals can be employed for unobtrusive surveillance, tracking suspicious movements in restricted areas or public spaces without compromising privacy [9].
4. Retail analytics: Retailers can use Wi-Fi signal-based recognition to analyze customer movements within stores, gaining insights into shopping patterns and improving store layouts for a better customer experience [10].

The CSI is crucial in Wi-Fi signal-based human action recognition due to its sensitivity to environmental changes caused by human movements. Namely, CSI captures subtle variations in signal properties as a person moves, enabling the development of accurate algorithms for action recognition. By examining phase shifts and amplitude changes, it becomes possible to distinguish between different actions, even those involving intricate gestures. Moreover, CSI provides a level of abstraction from the raw Wi-Fi signal, allowing for more robust and generalized models. This abstraction aids in the creation of machine learning models that can recognize actions across different Wi-Fi hardware and environmental conditions. In summary, Wi-Fi signal-based human action recognition represents a cutting-edge approach to understanding and interpreting human behavior without compromising privacy. The utilization of CSI enables the development of accurate and versatile models, paving the way for a wide array of applications in healthcare, smart homes, security, and retail analytics. As technology continues to advance, Wi-Fi signal-based human action recognition stands as a testament to the ingenuity of leveraging everyday technologies for innovative solutions.

Wi-Fi CSI-based human action recognition systems have gained significant attention in recent years due to their non-intrusive nature and wide range of applications, from smart homes to healthcare monitoring. As already mentioned, these systems leverage the subtle changes in Wi-Fi signals caused by human movements, which are then analyzed using deep learning algorithms to classify various human actions. Despite the promising advancements, a critical issue undermines the reliability and generalizability of these systems: data leakage.

Data leakage, a phenomenon where information from outside the training dataset improperly influences the model, can severely distort the performance metrics of machine learning models. In the context of Wi-Fi CSI-based human action recognition, data leakage often occurs due to the absence of the subject-based partitioning of datasets. The subject-based partitioning of a database in the context of machine learning refers to dividing the dataset based on distinct subjects or individuals. This method ensures that all data related to a specific subject are grouped together and used exclusively in one of the subsets (training, validation, or test set). For example, in a dataset containing activity data from multiple individuals, subject-based partitioning would involve assigning each person's entire set of data to either the training set, validation set, or test set, without mixing data from the same person across these subsets. This approach helps in evaluating the generalization capability of the model to new, unseen subjects, ensuring that the model is not merely learning individual-specific patterns but can generalize to different subjects. The abundance of subject-based partitioning involves splitting the dataset in a way that data from the same individuals appear in both training and testing sets, leading to artificially inflated accuracy scores. Such partitioning fails to reflect real-world scenarios where the system must generalize to recognize actions from unseen individuals. To illustrate this, let us consider a scenario where a company develops a health monitoring device in country *A* using data from local participants. Namely, the device's algorithm is trained to recognize various health metrics and behaviors using these data. However, this device is intended to be sold and used in country *B*, with a different population. The problems of partitioning without respect to subjects (or humans) are the following:

- If the data are randomly partitioned, data from the same individuals could end up in both the training and test sets.
- The model may learn specific characteristics of those individuals, leading to high accuracy during testing.

- However, this performance might not translate to new users in country *B*, as the model has not learned to generalize beyond the specific subjects in country *A*.

    In contrast, the advantages of subject-based partitioning of the database are the following:

- By using subject-based partitioning, the model is trained on one group of individuals and tested on a completely separate group.
- This ensures that the model learns to generalize patterns that apply broadly to different individuals, improving its performance when deployed in country *B*.

Based on the above considerations, we may say that partitioning without respect to humans can lead to data leakage, where information from the test set influences the training process, resulting in an overestimation of the model's performance. On the other hand, subject-based partitioning eliminates this risk by ensuring that no individual's data appear in both the training and test sets, providing a more accurate evaluation of the model's true performance.

This paper aims to systematically identify and analyze instances of data leakage in published Wi-Fi CSI-based human action recognition systems that utilize deep learning. By examining the dataset partitioning strategies employed, we demonstrate how subject-based partitioning can lead to misleading performance evaluations. Furthermore, we discuss the implications of these findings for the development of robust and reliable human action recognition systems and propose best practices to mitigate data leakage. Our study underscores the importance of rigorous dataset management and evaluation protocols to ensure the integrity and applicability of deep learning models in real-world applications. By addressing the issue of data leakage, we aim to pave the way for more accurate and dependable human action recognition systems that can truly leverage the potential of Wi-Fi CSI technology.

Contributions

This paper makes the following key contributions to the field of Wi-Fi CSI-based human action recognition.

- Identification and analysis of data leakage: We conduct an in-depth analysis of data leakage in three published Wi-Fi CSI-based human action recognition methods. Our study highlights how improper dataset partitioning, specifically, the failure to partition data with respect to individual subjects, can lead to artificially inflated performance metrics.
- Evaluation with proper dataset partitioning: We re-evaluate the aforementioned methods using subject-based partitioning strategies, demonstrating a significant decline in performance. This underscores the critical importance of proper dataset management in developing robust and generalizable models.
- Comparison of preprocessing techniques: Our study reveals that the impact of various preprocessing techniques, such as Canny, Sobel, Prewitt, and LoG filtering, is less significant when correct data partitioning is applied. This finding emphasizes the primacy of proper data splitting over preprocessing choices.
- Continuation of previous work: Building on our prior research where we analyzed data leakage in another published method [11], this paper extends our efforts to ensure the validity and reliability of Wi-Fi CSI-based human action recognition systems.

By addressing these issues, we aim to improve the methodological rigor in the evaluation of human action recognition systems and provide a foundation for future research to build more trustworthy and generalizable models.

2. Preliminaries on RSSI and CSI

In the digital age, connectivity is the backbone of our daily lives, seamlessly integrating into our routines. One of the most transformative innovations in this realm is Wi-Fi, a wireless communication technology that has revolutionized the way we access and share information. The received signal strength indicator (RSSI) is a measurement used in wireless

communication systems, including Wi-Fi, to quantify the power level of the received signal. RSSI represents the strength or intensity of the radio signal as it is received by a wireless device, such as a Wi-Fi-enabled device or router. It is usually expressed in decibels. The RSSI value provides an indication of the signal strength between a transmitting device (e.g., a Wi-Fi router) and a receiving device (e.g., a smartphone, laptop, or another Wi-Fi-enabled device). A higher RSSI value typically indicates a stronger and more robust signal, while a lower value suggests a weaker signal. It is important to note that RSSI alone may not provide a comprehensive assessment of the overall quality of a wireless connection. The signal quality can be affected by various factors, including interference, obstacles, and the distance between the devices. Additionally, different manufacturers and devices may use slightly different RSSI scales, making it advisable to interpret RSSI values in the context of a specific device or system. Formally, RSSI can be expressed as based on the received power (P_r)

$$RSSI = 10 \cdot \log_{10}(\frac{P_r}{P_0}), \tag{1}$$

where P_r is the received power in milliwatts (mW) and P_0 is a reference power, typically 1 mW. This formula can be simplified in practical applications where RSSI is measured directly by Wi-Fi hardware. In many cases, RSSI values are provided in dBm (decibels relative to one milliwatt), which inherently use a logarithmic scale.

CSI refers to a set of parameters that describe the characteristics of the wireless communication channel between a transmitter (e.g., Wi-Fi router) and a receiver (e.g., Wi-Fi-enabled device). CSI provides detailed information about how the radio signals traverse the wireless medium, including phase, amplitude, and frequency information. CSI is valuable for advanced signal processing techniques, beamforming, and fine-tuning of communication protocols to optimize wireless performance in various conditions. Unlike RSSI, which provides a general indication of signal strength, CSI offers a more detailed and nuanced view of the channel conditions. It is particularly useful for advanced applications in wireless communications, such as multiple-input multiple-output (MIMO) systems, beamforming, and other techniques that leverage the spatial and temporal characteristics of the channel to enhance data rates and reliability. CSI contains amplitude and phase measurements as

$$h = |h| \cdot e^{j \cdot \sin \theta}, \tag{2}$$

where $|h|$ and θ stand for the amplitude and the phase, respectively. In this paper, all subcarriers of CSI are considered for Wi-Fi-based human action recognition (HAR).

3. Related Works

The following section reviews existing literature pertinent to our study, divided into two key areas: human action recognition and data leakage. The first subsection provides an overview of the methodologies and advancements in human action recognition, particularly focusing on systems utilizing Wi-Fi CSI and deep learning. The second subsection explores the concept of data leakage in machine learning, examining its causes, effects, and the strategies proposed to detect and mitigate this issue across various domains.

3.1. Wi-Fi CSI-Based Human Action Recognition

In this subsection, a brief overview on existing CSI-based Wi-Fi HAR methods is given. Specifically, publicly available databases related to Wi-Fi-based HAR are outlined first. Subsequently, representative algorithms and architectures are discussed. Using Wi-Fi for HAR has a number of advantages, such as convenience, simplicity, privacy protection, and the low cost of Wi-Fi devices [12]. Further, human bodies can reflect Wi-Fi signals for HAR well enough, even for through-wall scenarios [13]. In particular, the vast majority of Wi-Fi-based methods for HAR have utilized CSI, since—as already mentioned—it is the fine-grained information calculated from the raw signal and Wi-Fi signals reflected from a moving person typically produce unique CSI on a receiver [14,15].

In 2018, Guo et al. [16] published the Wi-Fi–activity recognition (WiAR) database and the HuAc architecture, which is a combination of Wi-Fi-based and Kinect-based [17] activity recognition systems. Specifically, the WiAR database was collected in three different settings, i.e., empty room, meeting room with a desk, and office. Further, sixteen activities were recorded, such as horizontal arm wave, high arm wave, two hands wave, high throw, draw X, draw tick, toss paper, forward kick, side kick, bend, hand clap, walk, phone call, drink water, sit down, and squat. In this database, the authors published the raw RSSI and CSI signals related to each activity. A similar database was published by Wang et al. [18], who collected CSI data for HAR and indoor localization, containing six different activities, such as hand up, hand down, hand left, hand right, hand circle, and hand cross, which may have significance for human–computer interaction. In contrast, the publishers of StanWiFi dataset [6] asked six subjects to carry out six daily activities, i.e., sitting down, standing up, lying down, running, walking, and falling, in an indoor environment. In this environment, a transmitter (Wi-Fi router with one antenna) and a receiver (computer with NIC 5300 and three antennas) were placed 3 m from each other. Unlike the other publicly available datasets, Alazrai et al. [19] compiled Wi-Fi-based human-to-human interaction datasets, where pairs of subjects perform different actions, such as approaching, departing, handshaking, high fives, hugging, kicking with the left leg, kicking with the right leg, pointing with the left hand, pointing with the right hand, punching with the left hand, punching with the right hand, and pushing.

A time–frequency diagram, also known as a spectrogram, is a visual representation of how the frequency content of a signal changes over time [20]. It is a two-dimensional plot, where the x-axis represents time, the y-axis represents frequency, and the intensity of each point or pixel in the plot represents the magnitude or power of the signal's frequency components at that particular time and frequency. Using energy changes at different frequencies along time, time segments of actions and human actions can be detected and identified. For instance, the E-eyes system [14] utilizes a matching algorithm to compare the spectogram of a given CSI to already known profiles for human activity identification. In [21,22], researchers demonstrated the potential of Wi-Fi signals for the recognition of small scale human movements. Specifically, Wang et al. [21] introduced the WiHear framework, where fine-grained radio reflections from the mouth were detected and analyzed to reconstruct people's speech. Similarly, Tan et al. [22] utilized fine-grained CSI from a commodity Wi-Fi device, but the authors identified finger gestures, such as zoom out, zoom in, circle left, circle right, swipe left, swipe right, flip up, and flip down. In contrast, large-scale human movements were detected in the WiSee [23], WiTrack [24], and Wi-Vi [25] projects. To be more specific, Pu et al. [23] detected nine large-scale movements, i.e., push, dodge, strike, pull, drag, kick, circle, punch, and bowling, to create a Wi-Fi based human–computer interface. In the WiTrack [24] project, the focus was on the tracking of the human body and the human body parts using the Wi-Fi signal. In the Wi-Vi [25] project, researchers demonstrated that Wi-Fi signals enable the detection of moving objects through walls. The Wi-Sleep system [26] extracts rhythmic patterns from the CSI to monitor a person's sleep. It was demonstrated that movements during sleep, i.e., postures and rollovers, can be reliably detected. Chen et al. [27] utilized discrete wavelet decomposition to extract features from the CSI, and subsequently, a support vector machine (SVM) was trained to identify table tennis human actions. In [28], researchers applied the time variability and diversity of CSI to train a SVM [29] for human fall detection. Proposals for Wi-Fi-based human fall detection systems can also be found in [30–33]. The advent of deep learning has also changed the landscape in Wi-Fi-based HAR, recently. Zhou et al. [34] published a hybrid system, where features from CSI were extracted using a deep CNN, and human actions were classified with a trained SVM. In [35], a similar architecture was introduced for indoor fingerprinting. Namely, a deep network with four hidden layers was trained in an offline phase, and subsequently, a probabilistic method based on the radial basis function was used in an online phase for location estimation. Similarly, a fingerprinting system was discussed in [36], where a deep network with three hidden layers was trained on calibrated

phase data. In [37], a deep spare autoencoder was introduced, which learned discriminative features from CSI streams. Inspired by the success of UNet [38] in image segmentation, Wang et al. [39] implemented and introduced temporal UNet to learn a mapping from CSI data to action categories. In [40–42], long short-term memory (LSTM) networks [43] were employed to predict human actions from CSI. Specifically, Chen et al. [40] implemented an attention-based [44] bi-directional LSTM [45] to learn features from CSI in two directions. In contrast, Huang et al. [41] extracted features from CSI with a CNN and predicted human actions based on the extracted deep features and an LSTM. Similarly, Sheng et al. [42] built their system on the exploitation of deep features, but a bi-directional LSTM predicted human actions. Another line of works have focused on the fusion of Wi-Fi and other modalities for HAR. For instance, Zou et al. [46] implemented a two-stream CNN [47], which accepts CSI images at one stream and RGB images at the other stream, for HAR. In contrast, Memmesheimer et al. [48] took skeleton [49], inertial [50], and Wi-Fi signals and represented them as 2D images. Subsequently, these images were fed into a CNN to predict human actions.

A comprehensive overview on HAR is out of the scope of this paper. In [12], HAR algorithms were reviewed with respect to different data modalities, such as RGB image, skeleton, depth, infrared, point cloud, event stream, audio, acceleration, radar, and Wi-Fi. A similar but earlier study was written by Vrigkas et al. [51] In contrast, the studies of Zhang et al. [2], Pareek et al. [52], and Kong et al. [53] placed the focus on vision-based algorithms. In [54], Chen et al. summarized deep learning methods for sensor-based HAR. In [4], a state-of-the-art study was presented on HAR utilizing CSI. A similar study was published by Liu et al. [55], who reviewed the literature related to the wireless sensing of human activity.

3.2. Data Leakage in Machine Learning Research

The concept of data leakage in machine learning research refers to situations where information from outside the training dataset is inadvertently used to create the model, leading to overly optimistic performance estimates that do not generalize well to unseen data. This concern permeates various research domains, highlighting the complexity and multifaceted nature of preventing data leakage and ensuring robust, generalizable machine learning models. Data leakage poses several significant threats to the integrity and reliability of machine learning models [56]. Primarily, it can lead to overestimated performance metrics, giving a false sense of accuracy and robustness [57]. When a model inadvertently learns from data it should not have access to during training, it can fail to generalize to new, unseen data, thereby reducing its effectiveness in real-world applications. Additionally, data leakage can undermine the trust in and credibility of the research findings, potentially misleading stakeholders and decision-makers who rely on these models for critical applications [58]. In security-sensitive areas, such as healthcare or finance, data leakage can result in dire consequences, including compromised privacy and erroneous decision-making, further emphasizing the necessity of stringent measures to detect and prevent data leakage.

In the field of healthcare and particularly in cancer diagnosis, Samala et al. [59] underscore the hazards of data leakage when classifying breast cancer using deep neural networks, suggesting that validation set performance might be misleading compared to independent test set evaluations, possibly leading to overoptimistic assessments of a model's true predictive capability. Similar concerns were echoed by Rosenblatt et al. [60,61], who examined the effects of data leakage on connectome-based machine learning models used in neuroimaging data, indicating that prediction performance could be inflated through feature selection and repeated subject data, thus stressing the importance of avoiding data leakage to ensure model validity and reproducibility. Dong [62] highlighted a common source of leakage, where variables used to train the algorithm may inadvertently contain information about the outcome variable, resulting in unreliable and overly optimistic predictive performance. Addressing the prevention of data leakage, Moghaddam and

Zincir-Heywood [63] delved into the impact of stronger encryption algorithms on reducing data leakage in encrypted payloads analyzed using supervised machine learning, showing a technical approach to curtail data exposure risks. In [64], Poldrack et al. emphasized the distinction between correlation and prediction, highlighting that many neuroimaging studies erroneously claim predictive power based on correlations. It was pointed out that to establish evidence for prediction, models must be tested on independent data, not the same data used for parameter estimation. The authors recommended several best practices for predictive modeling: not reporting in-sample model fit as predictive accuracy, ensuring cross-validation includes all data operations, avoiding small samples, examining and reporting multiple accuracy measures, computing the coefficient of determination correctly, and using k-fold cross-validation over leave-one-out. Kapoor and Narayanan [65,66] introduced a taxonomy of eight types of leakage—i.e., not having a separate test set, preprocessing on the training and test sets, feature selection jointly on the training and test sets, duplicate data points, illegitimate features, temporal leakage, non-independence between the training and test sets, and sampling bias—ranging from textbook errors to open research problems. They proposed the implementation of "model info sheets" as a tool to help researchers identify and prevent leakage. These sheets consist of 21 questions that guide researchers in ensuring the legitimacy of their models and the correctness of their scientific claims.

4. Methods

In this section, we detail the methodologies employed in three different Wi-Fi CSI-based human action recognition systems, where data leakage was detected by us. Each subsection provides a comprehensive description of the respective method, focusing on the implementation details, dataset partitioning strategies, and the training parameters used. By examining these methods, we aim to illustrate how improper dataset partitioning, particularly the absence of subject-based splits, leads to significant data leakage and inflated performance metrics.

4.1. An Efficient Human Activity Recognition System Using Wi-Fi Channel State Information

Jiao and Zhang [67] presented a framework for HAR using Wi-Fi CSI in a paper entitled "An efficient human activity recognition system using Wi-Fi channel state information" and published in *IEEE Systems Journal*. The general workflow of this approach is depicted in Figure 1. The framework utilizes Gramian angular fields (GAFs) to convert CSI into images, which are then analyzed by a CNN to extract activity information. According to the reported results, the proposed framework achieves excellent performance (accuracy rate is 99.4% and F1 score is 99.4%), and it also exhibits low complexity compared to classical deep learning models.
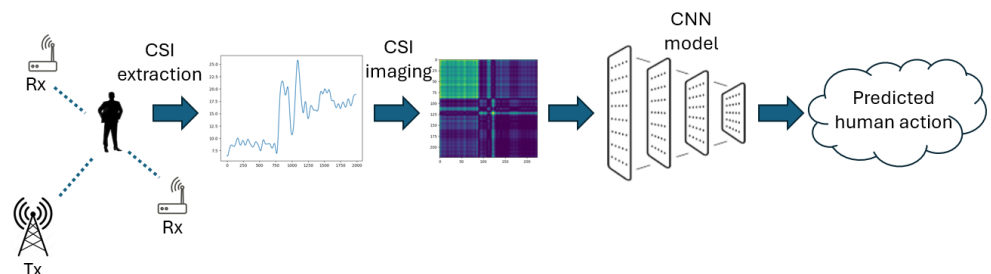
Figure 1. The general workflow of the method proposed by Jiao and Zhang [67]. First, CSI signals are converted into images, which are then analyzed by a CNN to predict human actions, using Gramian angular fields.

Since CSI data are given as time series of complex numbers, the goal of the CSI imaging module is converting these time series to two-dimensional images, which can later be used for fine-tuning CNN and transformer architectures. To this end, Gramian angular fields

(GAFs) are applied, which are methods for transforming time series data into a matrix format that can be used for various purposes, such as visualization, feature extraction, or classification. Since CSI contains complex numbers, CSI amplitude is considered in the imaging process. In the following, the amplitude values are meant if the term CSI data is used.

The GAF is constructed as follows [68]. Let us suppose that we are given time series $X = \{x_1, x_2, \ldots, x_n\}$ of length n. The first step is the normalization of the time series to the $[-1; 1]$ range:

$$\tilde{x}_i = \frac{x_i - \max(X) + (x_i - \min(X))}{\max(X) - \min(X)}. \tag{3}$$

Subsequently, the time series is converted to polar coordinates:

$$\begin{cases} \phi_i = \arccos(\tilde{x}_i), \ \tilde{x}_i \in \tilde{X}, \\ r_i = \frac{t_i}{N}, \ t_i \in \mathbf{N}, \end{cases} \tag{4}$$

where t_i and N stand for the time step and the normalization factor, respectively. Further, N is set to 1 according to [69]. Finally, $GASF$ and $GADF$ can be defined as follows:

$$GASF = \cos(\phi_i + \phi_j), \tag{5}$$

$$GASF = \begin{bmatrix} \cos(\phi_1 + \phi_1) & \ldots & \cos(\phi_1 + \phi_N) \\ \cos(\phi_2 + \phi_1) & \ldots & \cos(\phi_2 + \phi_N) \\ \vdots & \ddots & \vdots \\ \cos(\phi_N + \phi_1) & \ldots & \cos(\phi_N + \phi_N) \end{bmatrix}, \tag{6}$$

$$GADF = \sin(\phi_i - \phi_j), \tag{7}$$

$$GADF = \begin{bmatrix} \sin(\phi_1 - \phi_1) & \ldots & \sin(\phi_1 - \phi_N) \\ \sin(\phi_2 - \phi_1) & \ldots & \sin(\phi_2 - \phi_N) \\ \vdots & \ddots & \vdots \\ \sin(\phi_N - \phi_1) & \ldots & \sin(\phi_N - \phi_N) \end{bmatrix}. \tag{8}$$

Figure 2 illustrates the conversion of an example time series into GASF and GADF. As described above, the time series is first normalized. Next, the normalized signal is converted into polar coordinates. Finally, GASF and GADF can be obtained from the polar coordinates using Equations (6) and (8), respectively.

$GASF$ and $GADF$ can also be expressed via matrix operations:

$$GASF = \tilde{X}^T\tilde{X} - \sqrt{I - \tilde{X}^2}^T\sqrt{I - \tilde{X}^2}, \tag{9}$$

$$GADF = \sqrt{I - \tilde{X}^2}^T\tilde{X} - \tilde{X}^T\sqrt{I - \tilde{X}^2}, \tag{10}$$

where X is the given time series, and $\tilde{X}$ is the normalized time series obtained from X using Equation (3). Further, I stands for a vector containing only ones. According to the authors' results, GADF slightly outperforms GASF. This is why we opted to use GADF in our reimplementation and experiments. Figure 3 illustrates the process of CSI imaging through a series of subplots. The first subplot shows the raw CSI amplitude, capturing the initial, unprocessed signal data. The second subplot presents the filtered CSI signal, where noise and irrelevant variations have been reduced to highlight significant patterns. The final subplots depict the GASF and GADF, which transform the filtered CSI signal into visual representations suitable for 2D deep learning model input. These steps collectively demonstrate the transformation of raw CSI data into structured forms that enhance the capability of human action recognition systems.
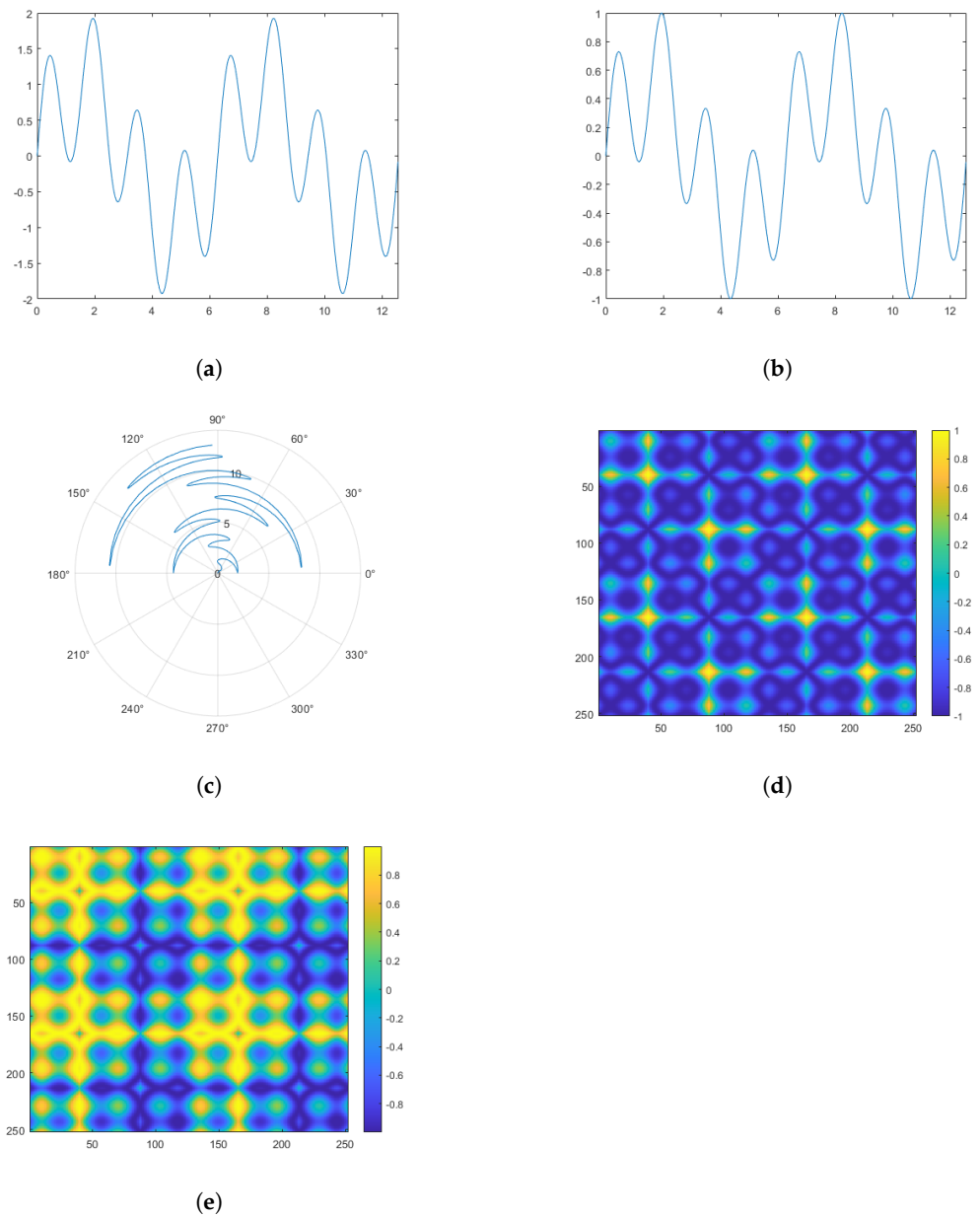
Figure 2. Illustration of GASF and GADF computation. (**a**) Example time series. (**b**) Normalized time signal obtained using Equation (3). (**c**) Mapping to polar coordinates using Equation (4). (**d**) GASF obtained using Equation (6). (**e**) GADF obtained using Equation (8).

As already mentioned, the proposed CNN analyzes the CSI images created with GADF to obtain human activity information. The structure of the CNN is depicted in Figure 4. Each convolutional layer has a kernel size of 3×3 and a varying number of channels (64, 128, 256, and 512). Batch normalization layers are employed after each convolutional layer to increase convergence speed, followed by rectified linear units (ReLUs) as activation functions. Additionally, the CNN utilizes maximum pooling and adaptive average pooling layers. Finally, a classifier with a dropout layer, a linear layer, and a ReLU layer follows the adaptive average pooling to obtain the activity information from the CSI images. The hyperparameters used in this method are detailed in Table 1. Besides the proposed CNN architecture, the authors fine-tuned three on ImageNet [70] database pretrained CNNs, such as ResNet50 [71], VGG19 [72], and ShuffleNet [73], on the Wi-Fi CSI databases for HAR. Although the process of fine-tuning was not explicitly described in the original paper, we can reasonably assume that a standard fine-tuning procedure was employed.

Namely, the final output layers of the pretrained networks were replaced to match the number of classes in the new task. For example, if the pretrained model was designed for 1000 classes—as in the case of ImageNet database—and the new task has 10 classes, the final layer was modified to output 10 classes. Next, the modified model was trained using the new dataset, and performance was monitored on a validation set to avoid overfitting. The hyperparameters of fine-tuning are identical to those used in the training of the proposed CNN structure and already given in Table 1.

(a)

(b)

(c)

(d)
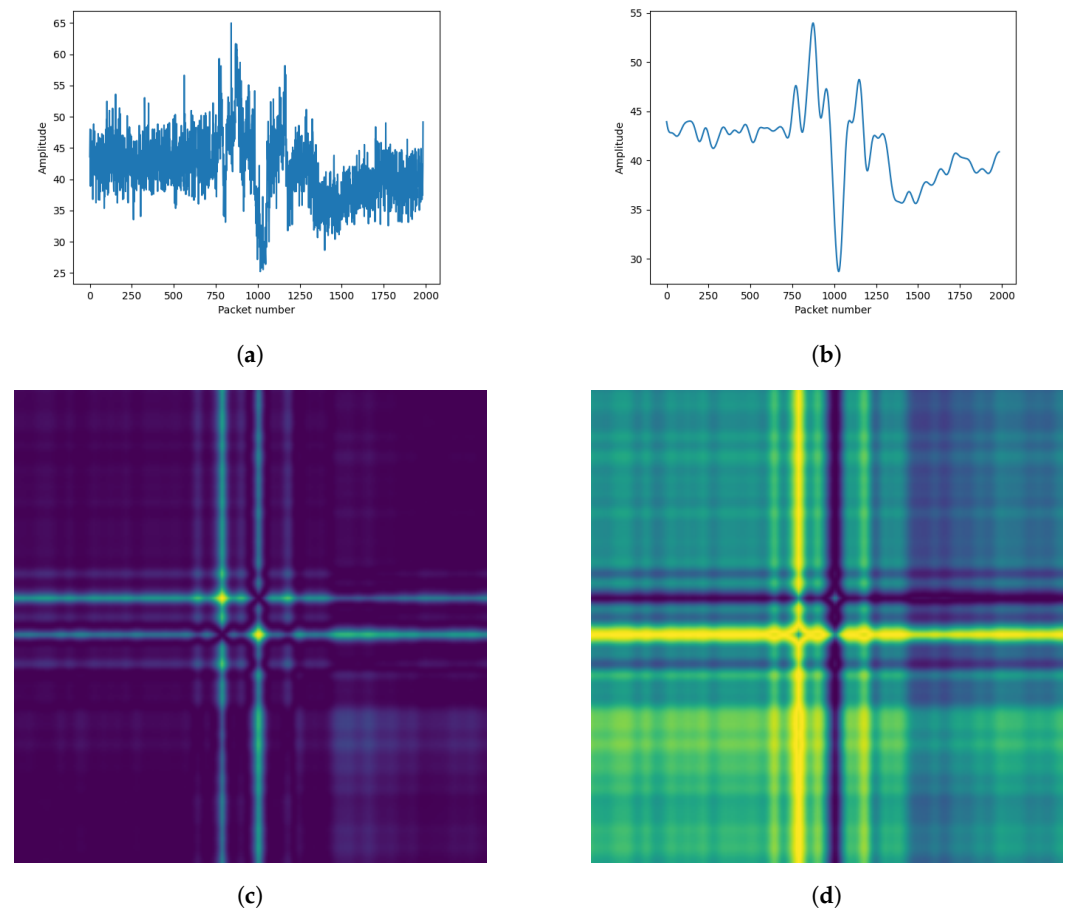
Figure 3. Illustration of CSI signal conversion to RGB image applied in the method of Jiao and Zhang [67]. (**a**) Raw CSI signal. (**b**) Filtered CSI signal. (**c**) GASF. (**d**) GADF.

Table 1. Hyperparameters used for the training of the CNN proposed by Jiao and Zhang [67].

| Parameter	Value
Loss function	Cross-entropy
Optimizer	Adam [74] ($\beta_1 = 0.9$, $\beta_2 = 0.99$, $\epsilon_2 = 1 \times 10^{-9}$)
Learning rate	0.001
Decay rate	0.8
Batch size	128
Dropout rate	0.5
Epochs	20
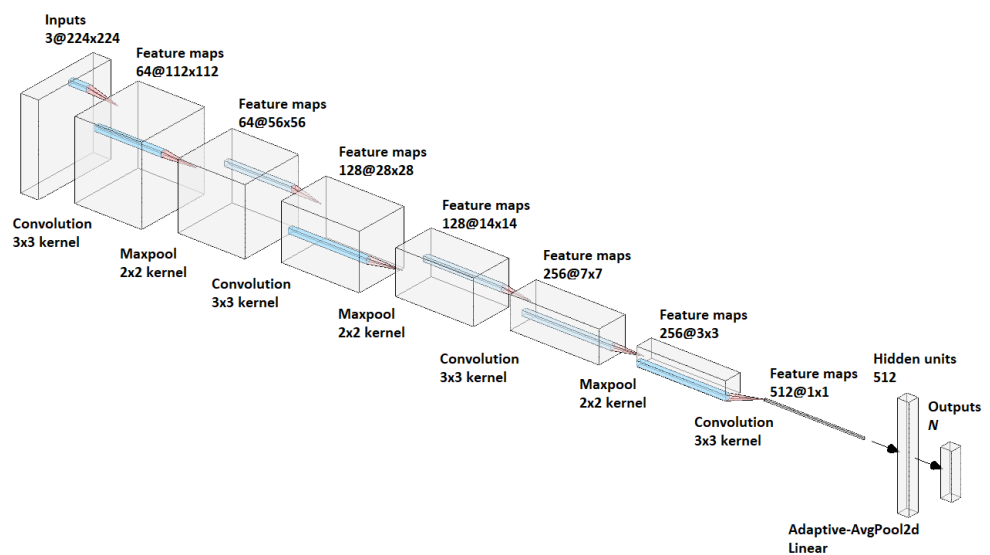
Figure 4. Structure of the CNN proposed and implemented by Jiao and Zhang [67] for Wi-Fi CSI-based HAR. Batch normalization layers were implemented after each convolutional layer to increase convergence speed, followed by ReLU as activation functions.

### 4.2. Human Activity and Gesture Recognition Based on Wi-Fi Using Deep Convolutional Neural Networks

The paper entitled "Human activity and gesture recognition based on Wi-Fi using deep convolutional neural networks" published in *Iraqi Journal for Electrical and Electronic Engineering* [75] proposed a similar Wi-Fi CSI-based HAR approach to those published by Jiao and Zhang [67] and discussed in the previous subsection. The general overview of the authors' method is depicted in Figure 5. The main difference lie in the creation of CSI images and the applied deep architectures. Namely, Jawad and Alaziz [75] employed deep learning models such as AlexNet [70], VGG19 [72], and SqueezeNet [76] for classification and feature extraction. Initially, outliers are removed from the amplitude of CSI stream using the Hampel filter algorithm. This is essential, as noise and outliers can impact the classification outcomes. The potential causes of noise and outliers include furniture, interference from neighboring devices, and transmitter transmitting power adaption. As already mentioned, the preprocessing stage involves using the Hampel algorithm to remove abnormal points from the amplitude of the CSI. This algorithm identifies and eliminates outliers based on a specified range. After preprocessing, the CSI data are converted to RGB images. Unlike in [67,77], Jawad and Alaziz [75] did not utilize time series imaging techniques, such as GADF, GASF, or recurrence plot transformation [78] but created an image from 30 subcarriers of CSI using MATLAB's *imagesc* function [79]—as illustrated in Figure 6. Data augmentation techniques were also implemented to reduce overfitting in deep learning models. The data augmentation involves introducing additional, slightly altered copies of the existing data or synthesizing new data to increase the amount of data available for training the models. Specifically, Jawad and Alaziz [75] applied random rotation in the range $[-30°, +30°]$, horizontal reflection, and random translation. Finally, several on ImageNet [80] database trained CNNs were fine-tuned for HAR-based Wi-Fi CSI. The parameters used for the fine-tuning of CNNs are shown in Table 2. The process of fine-tuning was as follows. First, the final output layers of the pretrained networks were replaced to match the number of classes in the new task. For example, if the pretrained model was designed for 1000 classes—as in the case of ImageNet database—and the new task had 10 classes, the final layer was modified to output 10 classes. The hyperparameters of fine-tuning are given in Table 2. The authors of this method applied a special evaluation technique, wherein the dataset was split equally into a training set (50% of images) and a validation set (50% of images). The performance results of the models were then assessed using validation accuracies, which were reported as the evaluation metrics.
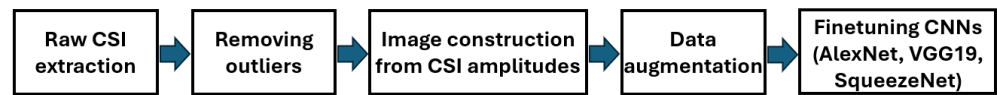
Figure 5. The general workflow of the Wi-Fi CSI-based HAR method proposed by Jawad and Alaziz [75].

Table 2. Hyperparameters used for the training (fine-tuning) of various on ImageNet [80] pre-trained CNNs.

	AlexNet [70]	**VGG19 [72]**	**SqueezeNet [76]**
Input image size	$227 \times 227 \times 3$	$224 \times 224 \times 3$	$227 \times 227 \times 3$
Loss function	cross-entropy	cross-entropy	cross-entropy
Learning rate	0.0001	0.0002	0.0002
Batch size	10	10	10
Epochs	20	20	30
Optimizer	SGDM	SGDM	SGDM
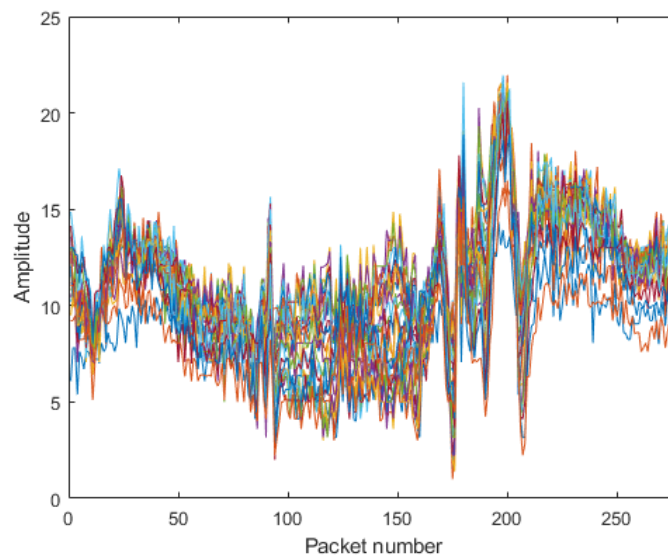
(**a**)
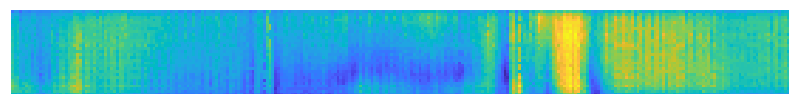
(**b**)

Figure 6. Illustration of CSI signal conversion to RGB image applied in Jawad et al.'s [75] method: (**a**) 30 Hampel filtered CSI signals. (**b**) CSI signals converted to RGB image using MATLAB's *imagesc* function.

4.3. Enhancing CSI-Based Human Activity Recognition by Edge Detection Techniques

In [81], Shahverdi et al. also utilized CSI data converted into RGB images. Prior to the conversion to RGB images, principal component analysis (PCA) [82], normalization, and linear discriminant analysis (LDA) [83] were applied to the raw CSI to reduce dimensionality and decrease noise. The authors' main contribution was the application of edge detection techniques on the RGB images as a preproceesing phase to improve the accuracy of activity recognition. The study utilized a CNN as a classifier, whose structure is depicted in Figure 7.
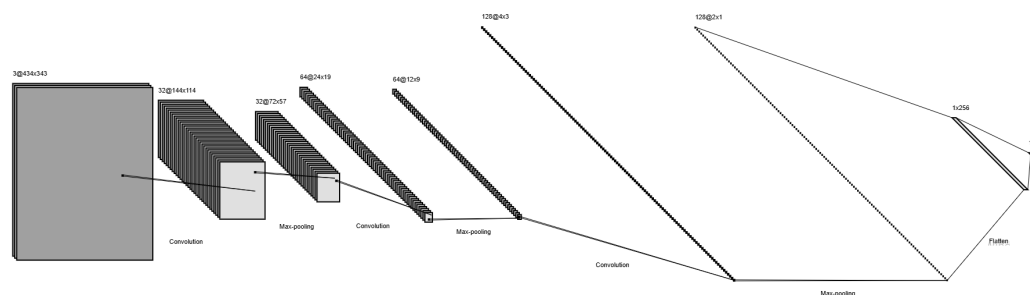
Figure 7. Structure of the CNN proposed and implemented by Shahverdi et al. [81] for Wi-Fi CSI-based HAR. To avoid overfitting, the authors implemented dropout with $p = 0.25$ parameter [84] after each convolutional and dense layer. Further, batch normalization layers were also implemented after each convolutional layer to further reduce overfitting. The authors used leaky ReLU [85] as the activation function.

Unfortunately, Shahverdi et al. [81] did not disclose the parameters used for the training of their model, such as the optimizer or the learning rate. To maintain consistency and ensure a fair comparison, we adopted the training parameters of Jiao and Zhang [67] (with the exception of the number of epochs, which was set to 30 to achieve better performance)—already given in Table 1—who implemented a similar CNN. This approach allowed us to standardize the evaluation process and provide a reliable assessment of the impact of data leakage on the performance of Wi-Fi CSI-based human action recognition systems.

4.4. Detected Data Leakage

Our analysis reveals a critical flaw in the data partitioning strategies employed by several Wi-Fi CSI-based human action recognition systems. Specifically, these methods partition the dataset without regard to individual subjects, leading to substantial data leakage. This practice involves splitting the data into training and testing sets in a way that allows data from the same individuals to appear in both sets.

Such an approach results in the model learning individual-specific features rather than generalized patterns of human actions, thereby inflating the performance metrics. When the same subjects are present in both the training and testing phases, the model can achieve unnaturally high precision rates, often exceeding 95%. This is because the model effectively memorizes the actions of specific individuals rather than learning to recognize actions across different individuals.

To illustrate the impact of this flaw, we re-evaluated the systems using a subject-based partitioning strategy, where data from certain individuals are entirely excluded from the training set and only used for testing. This approach is more reflective of real-world scenarios, where the system must generalize to recognize actions performed by unseen individuals. Our results demonstrate a significant drop in performance when subject-based partitioning is applied, highlighting the extent of data leakage in the original evaluations.

By ensuring that data from the same individuals are not simultaneously present in both training and testing sets, we can better assess the true generalizability of human action recognition models. This methodological correction is crucial for developing reliable and robust Wi-Fi CSI-based systems capable of performing accurately in diverse real-world applications.

5. Results

This section presents the results of our experiments, highlighting the impact of data leakage on the performance of Wi-Fi CSI-based human action recognition systems. We evaluate each method by retraining the models with and without subject-based dataset partitioning. The subsections provide a detailed comparison of the performance metrics, demonstrating how the absence of subject-based splits leads to artificially high precision rates due to data leakage. By contrasting these results, we underscore the importance of

rigorous dataset management to ensure the reliability and generalizability of these systems. The considered methods are reimplemented in Python 3.12.3 programming language using the PyTorch [86] deep learning library. Further, these methods are trained with the help of a GPU (graphics processing unit) server containing eight NVidia GeForce RTX 3090 [87] cards.

5.1. Results of the Reimplementation of "An Efficient Human Activity Recognition System Using Wi-Fi Channel State Information"

In this subsection, we compare the performance of Wi-Fi CSI-based HAR system proposed by Jiao and Zhang [67] when retrained using two different dataset partitioning strategies: with and without respect to individual subjects. To provide context for our analysis, we first summarize the key characteristics of the databases in Table 3 used in the original paper [67] and in this study. It can be seen in this table that WiAR [88] contains CSI data grouped into 16 action categories, while the number of action categories is 6 in Widar3.0 [89]. The original dataset partitioning strategy employed by the examined system of Jiao and Zhang [67] allocated 70% of the data for training, 15% for validation, and 15% for testing. However, to ensure that each partition contained an integer number of subjects in the case of retraining with respect to humans, we adjusted the partitioning strategy to 60% for training, 20% for validation, and 20% for testing. This adjustment was necessary to maintain the integrity of subject-based partitioning and avoid fractional allocations of individual subjects because WiAR contains data from 10 volunteers, while Widar3.0 consists of data from 5 volunteers.

Table 3. Dataset's details of WiAR [88] and Widar3.0 [89].

Dataset Name	Action Labels	Dataset Size
WiAR [88]	two hands wave, high throw, horizontal arm wave, draw tick, toss paper, walk, side kick, bend, forward kick, drink water, sit down, draw X, phone call, hand clap, high arm wave, squat	62,415 images
Widar3.0 [89]	push, sweep, clap, slide, draw-Z, draw-N	80,000 images

The results from retraining with respect to humans and without respect to humans on WiAR [88] database are presented in Table 4. From these results, it can be seen that the obtained values from retraining with respect to humans are lower than those reported in the original paper. We suspect that some kind of data augmentation technique was not disclosed in the original study. Specifically, for the ResNet50 model, we can observe a significant difference. We found that we could increase the performance of the ResNet50 model by increasing the number of epochs. However, we chose not to do this because we wanted to adhere to the number of epochs used in the original study. In the case of retraining with respect to humans, the performance drop to 20–25% of the reported metrics. While this can be discouraging, it is essential to use the correct data partitioning method, which is with respect to humans. The results from retraining with respect to humans and without respect to humans on the Widar3.0 [89] database are presented in Table 5. In the case of Widar3.0 [89], we were able to reproduce the reported results with good accuracy. However, when applying the correct data partitioning with respect to humans, the performance drops to 39–45% of the reported metrics. The higher performance compared to the previous database can be attributed to the lower number of actions (5 vs. 16) and the potentially higher distinctness of these actions.

Table 4. Comparison of results on WiAR.

	Reported in [67]		Retrained w/o.r.t. Humans		Retrained w.r.t. Humans	
Architecture	Acc.	F1	Acc.	F1	Acc.	F1
ResNet50	0.994	0.994	0.801	0.800	0.204	0.195
VGG19	0.993	0.994	0.932	0.931	0.229	0.218
ShuffleNet	0.992	0.992	0.933	0.932	0.225	0.216
Proposed CNN	0.994	0.994	0.955	0.955	0.231	0.220

Table 5. Comparison of results on Widar3.0.

	Reported in [67]		Retrained w/o.r.t. Humans		Retrained w.r.t. Humans	
Architecture	Acc.	F1	Acc.	F1	Acc.	F1
ResNet50	0.993	0.993	0.990	0.990	0.390	0.382
VGG19	0.992	0.993	0.991	0.991	0.444	0.440
ShuffleNet	0.991	0.991	0.991	0.991	0.443	0.439
Proposed CNN	0.993	0.993	0.992	0.993	0.456	0.440

Figures 8 and 9 depict the training curves without respect to humans and with respect to humans, respectively. This figure provides insightful conclusions. When the data split is performed without respect to humans, there is a strong correlation between training and validation accuracy. The validation accuracy closely follows the trends of the training accuracy, with only a slight difference. This alignment indicates that the model is effectively learning from the training data and generalizing well to the unseen validation data. On the other hand, when the data are split according to human subjects, a noticeable difference arises between the training and validation accuracy. While the training accuracy steadily improves, validation accuracy plateaus, suggesting that the model struggles to generalize effectively to unseen data.
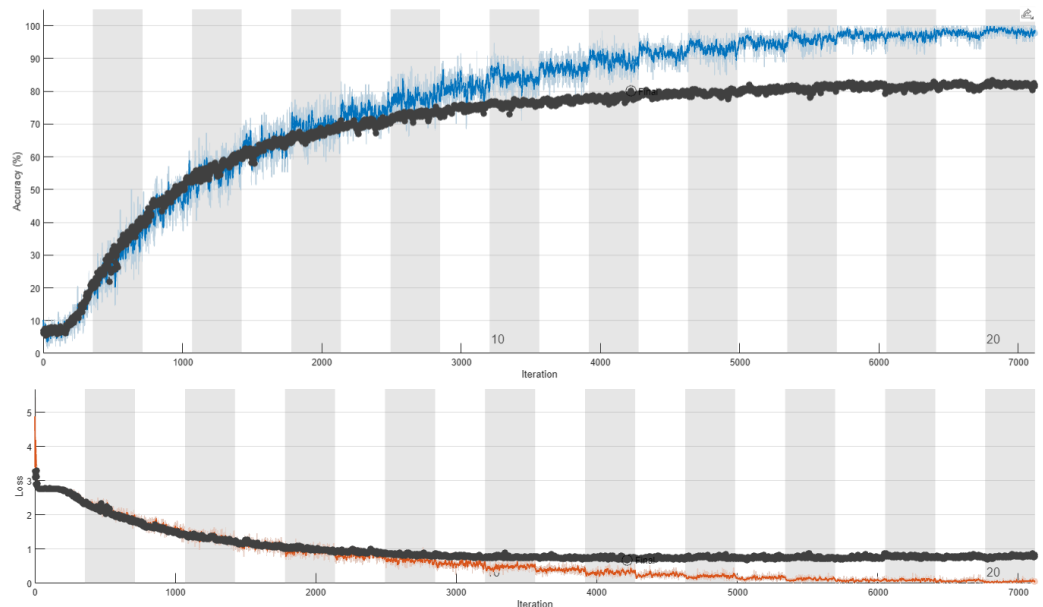
Figure 8. Training of ResNet50 on WiAR [88] without respect to humans. In the upper figure, the training accuracy is represented by the blue line, whereas the validation accuracy is depicted in black. In the lower figure, the training loss is indicated in red, and the validation loss is shown in black.
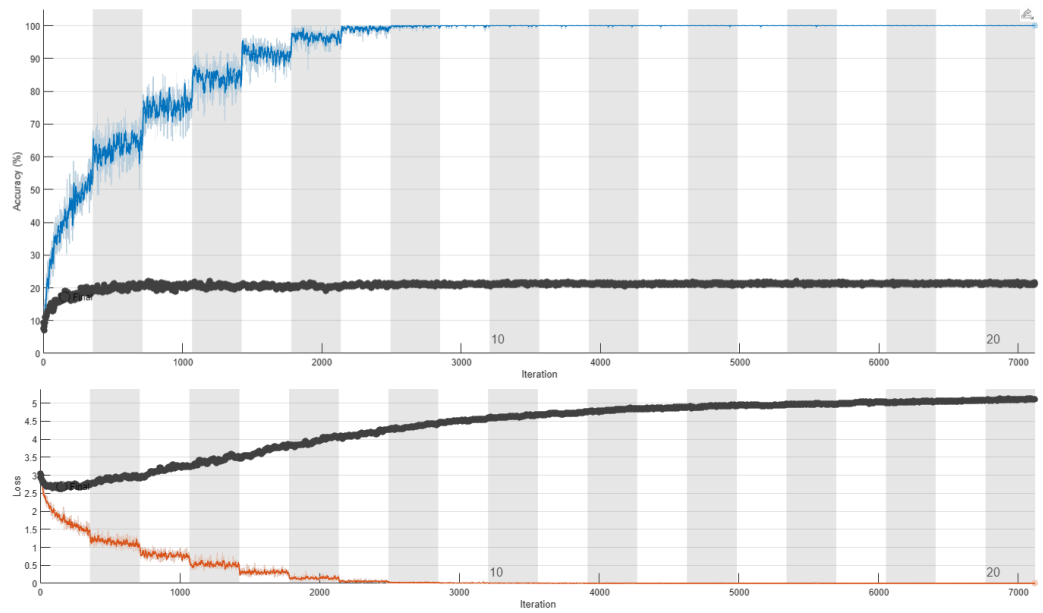
Figure 9. Training of ResNet50 on WiAR [88] with respect to humans. In the upper figure, the training accuracy is represented by the blue line, whereas the validation accuracy is depicted in black. In the lower figure, the training loss is indicated in red, and the validation loss is shown in black.

The confusion matrices of fine-tuned ResNet50 in Figure 10 illustrate the results on the WiAR test set from retraining without respect to humans and with respect to humans. It can be clearly seen that the results of retraining with respect to humans are significantly less favorable compared to those of retraining without respect to humans. The diagonal elements of the confusion matrix illustrating the results of the data split without respect to humans exhibit less variation than those of the data split with respect to humans, where the diagonal elements vary in the range of 2.5–56.4%.

(a)

Figure 10. *Cont.*

True \ Predicted	bend	draw X	draw tick	drink water	forward kick	hand clap	high arm wave	high throw	horizontal arm wave	phone call	side kick	sit down	squat	toss paper	two hands wave	walk
bend	8.1%	7.8%	1.4%	21.4%	12.8%	1.9%	0.8%	1.4%		5.6%	16.9%	0.6%	15.3%	2.2%	0.3%	3.6%
draw X	6.1%	26.9%	1.1%	7.8%	2.8%	0.3%	3.9%	0.6%	0.6%	3.1%	5.0%	2.8%	36.9%	1.1%	0.3%	0.8%
draw tick	1.9%	3.6%	2.5%	8.6%	2.5%	3.3%	0.6%	1.7%	2.2%	23.6%	21.7%	0.8%	1.9%	7.2%	8.1%	9.7%
drink water	6.7%	16.4%	4.2%	14.2%	9.2%	1.9%	2.2%	2.8%		8.1%	9.2%	0.3%	16.7%	2.2%	2.2%	3.9%
forward kick	13.3%	6.9%	1.4%	26.1%	16.4%	1.7%	0.8%	0.6%		4.2%	6.4%	0.3%	17.8%	0.6%	0.3%	3.3%
hand clap	0.3%	4.7%	3.3%	6.4%	0.3%	9.4%	2.5%	3.6%	1.9%	25.3%	7.8%	2.8%	2.5%	7.5%	10.3%	11.4%
high arm wave	1.9%	6.7%	1.4%	10.3%	3.9%	0.6%	7.8%	3.1%	4.7%	8.1%	11.9%	3.3%	25.3%	3.1%	5.8%	2.2%
high throw	0.8%	1.9%	1.4%	3.6%		1.4%	6.7%	33.9%	10.3%	16.9%	1.7%	1.4%	2.5%	4.4%	9.2%	3.9%
horizontal arm wave		0.8%	2.2%			0.8%	14.2%	19.4%	31.1%	10.8%	1.7%	6.4%	0.8%	6.1%	3.3%	2.2%
phone call	0.8%	2.8%	5.6%	4.7%	0.3%	5.3%		4.7%	3.3%	36.1%	3.6%	5.0%	0.3%	5.0%	13.9%	8.6%
side kick	3.6%	15.8%	3.3%	2.8%	1.4%	2.8%	2.8%	3.6%	0.3%	18.3%	18.1%	1.4%	2.5%	9.4%	8.3%	5.6%
sit down	1.1%	11.9%	2.5%	5.0%	0.6%	2.8%	13.3%	8.9%	15.0%	8.1%	2.8%	6.9%	11.4%	4.2%	2.5%	3.1%
squat	5.6%	14.4%	0.6%	5.6%	4.2%	0.3%	6.7%	1.4%	1.7%	0.3%	0.3%	1.1%	56.4%	0.6%	0.3%	0.8%
toss paper	4.7%	4.2%	4.4%	7.5%	2.2%	3.9%	3.1%	4.2%	11.4%	15.3%	11.7%	2.5%	8.3%	3.3%	6.9%	6.4%
two hands wave	0.8%	2.5%	1.9%	2.8%		2.8%	1.4%	19.4%	17.8%	18.9%	3.1%	0.3%	1.9%	3.6%	16.1%	6.7%
walk	2.2%	4.2%	2.8%	4.2%	1.1%	1.9%	6.1%	2.8%	9.7%	18.1%	10.6%	3.1%	3.3%	3.9%	7.8%	18.3%

(**b**)

Figure 10. Confusion matrices of fine-tuned ResNet50 obtained on WiAR [88] test set. (**a**) Results obtained from retraining without respect to humans. (**b**) Results obtained from retraining with respect to humans.

5.2. Results of the Reimplementation of "Human Activity and Gesture Recognition Based on Wi-Fi Using Deep Convolutional Neural Networks"

Similarly to the previously discussed paper, Jawad et al. [75] fine-tuned also several on ImageNet [80] database pretrained CNNs on the WiAR database. As already mentioned, the main difference lay in the CSI imaging method. The results reported by Jawad et al. [75] and the results coming from our own reimplementation are summarized in Table 6. From these results, it can be clearly seen that we were able to achieve the authors' results with relatively good approximation when data split without respect to humans was applied. Similarly to the previously discussed method, the performance drops to approximately 20% of the reported performance if the correct data split—with respect to humans—is applied.

Table 6. Comparison of results on WiAR.

|	Reported in [75]	Retrained without Respec to Humans	Retrained with Respect to Humans
Architecture	Acc.	Acc.	Acc.
AlexNet	0.9917	0.97	0.215
SqueezeNet	1.0	0.98	0.215
VGG19	0.9625	0.94	0.201

5.3. Results of the Reimplementation of "Enhancing CSI-Based Human Activity Recognition by Edge Detection Techniques"

Shahverdi et al. [81] utilized the database of Moshiri et al. [90] for testing their HAR framework based on Wi-Fi CSI signals. Specifically, Moshiri et al. [90] applied the Nexmon Tool [91] and collected CSI data for seven daily human activities, such as walk, run, fall, lie down, sit down, stand up, and bend. The collected CSI matrices have 52 columns corresponding to the number of subcarriers and 600–1100 rows depending on the duration of individual activities. Further, it is important to note that each activity of this dataset was carried out 20 times by three users of different ages. The major details of the CSI-HAR [90] database are summarized in Table 7. Several CSI matrices illustrated as RGB images can be seen in Figure 11 as illustration. Furthermore, it is presumed that Shahverdi et al.'s [81] method utilized a 80%/20% split for training and testing, respectively, although the original

article unfortunately does not explicitly confirm this. In the case of training with respect to humans, we allocated two individuals to the training set and one individual to the test set, because—as already mentioned—the number of volunteers in CSI-HAR database [90] is three. This partitioning ensured that the training and testing sets were mutually exclusive with respect to the subjects involved.

Table 7. Dataset's details of CSI-HAR [90].

| Dataset Name	Action Labels	Dataset Size
CSI-HAR [90]	bend, fall, lie down, run, sit down, stand up, walk	420 images

(a)

(b)

(c)

(d)
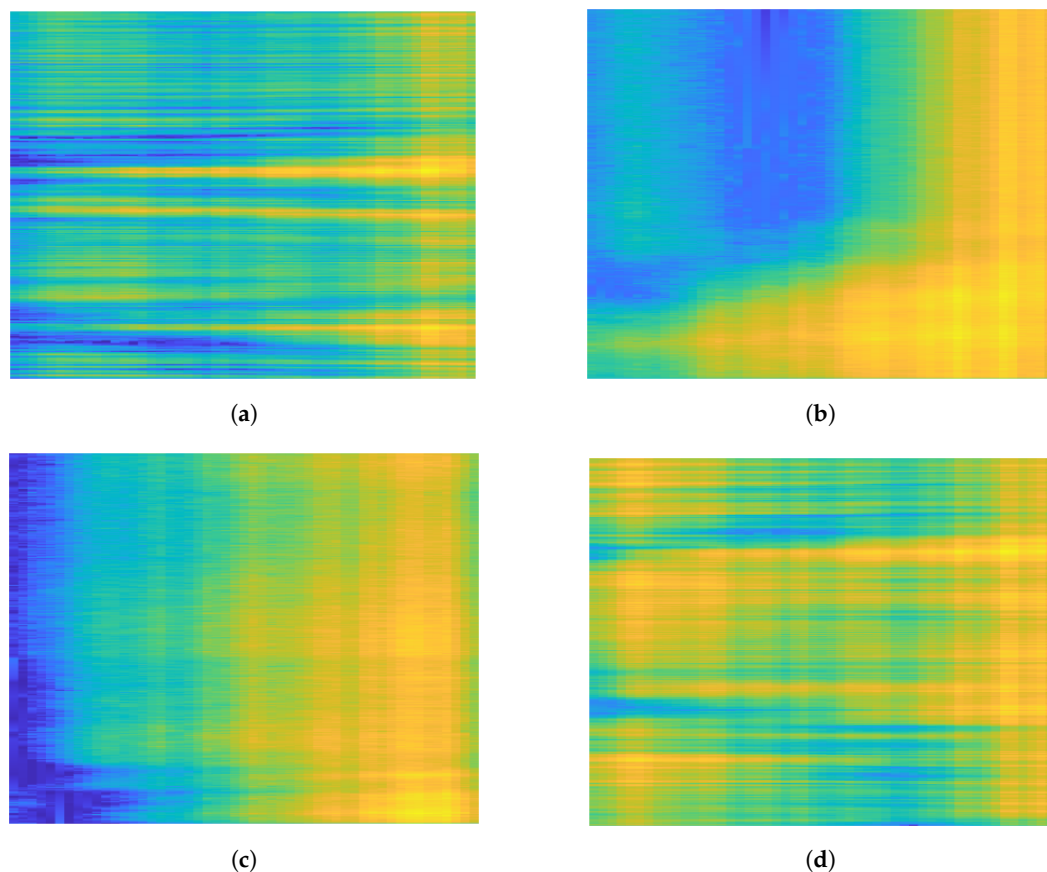
Figure 11. Illustration of RGB images in the CSI-HAR database [90]. (**a**) Run. (**b**) Sit down. (**c**) Stand up. (**d**) Walk.

The results are summarized in Table 8, which presents the outcomes from the original paper alongside the results from our retraining with and without respect to individual subjects. The results are summarized in Table 8. As it can be seen, the results obtained from retraining without respect to individual subjects are slightly lower than those reported in the original study. This discrepancy could be attributed to potential techniques, such as data augmentation, K-fold cross-validation, or regularization, which may have been employed but were not explicitly mentioned in the original paper. Despite this difference, the presence of data leakage was clearly demonstrated through this analysis. Retraining with respect to individual subjects yielded results that were approximately two-thirds of those reported in the original study. While this may seem discouraging, it underscores the importance of using subject-based data splits, which is the correct and more rigorous approach to ensure validity and generalizability of the model. Another important conclusion that can be drawn from our experimental results presented in Table 8 is that the effect of preprocessing techniques, such as Canny, Sobel, Prewitt, and LoG filtering, on the CSI matrices is less significant when the correct data split—with respect to individual subjects—is employed.

This also underscores that proper data partitioning plays a crucial role in achieving reliable results. In Figure 12, the corresponding confusion matrices are depicted. It can be observed that the diagonal elements show significantly greater variability in the case of data split with respect to humans. Namely, the values vary from 5% to 85%. Moreover, the fall activity is confused with run and stand up activities in half the cases. A similar conclusion can be drawn for the sit-down activity, which is often confused with the lie-down activity. Further, the run activity is confused with the walk activity with an extremely high probability (95%). In summary, we can declare that the results are significantly less favorable if the correct data split—with respect to humans—is applied, and several activities seem to be very challenging for deep architectures to recognize correctly based on Wi-Fi CSI signals. As a consequence, there is still a lot of space for future research. Figures 13 and 14 depict the training curves of the two different data-split strategies.
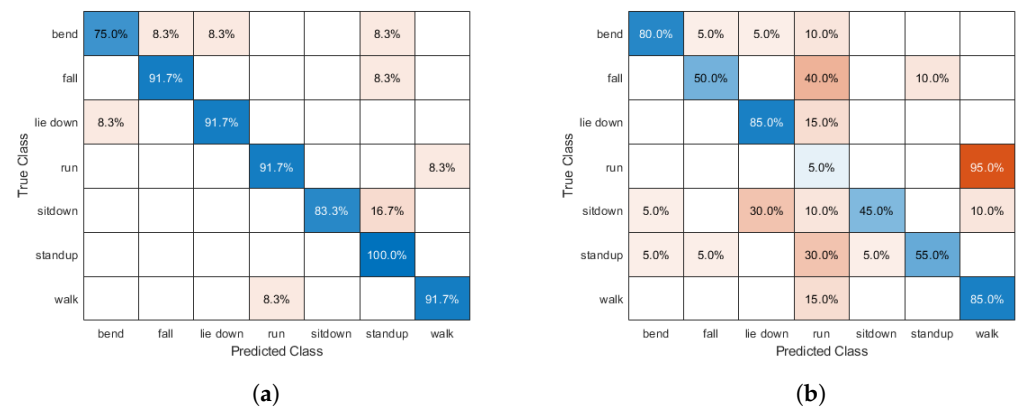

Figure 12. Confusion matrices of the deep architecture (RGB CSI images as input) proposed by Shahverdi et al. [81] obtained on CSI-HAR [90] test set. (**a**) Results obtained from retraining without respect to humans. (**b**) Results obtained from retraining with respect to humans.
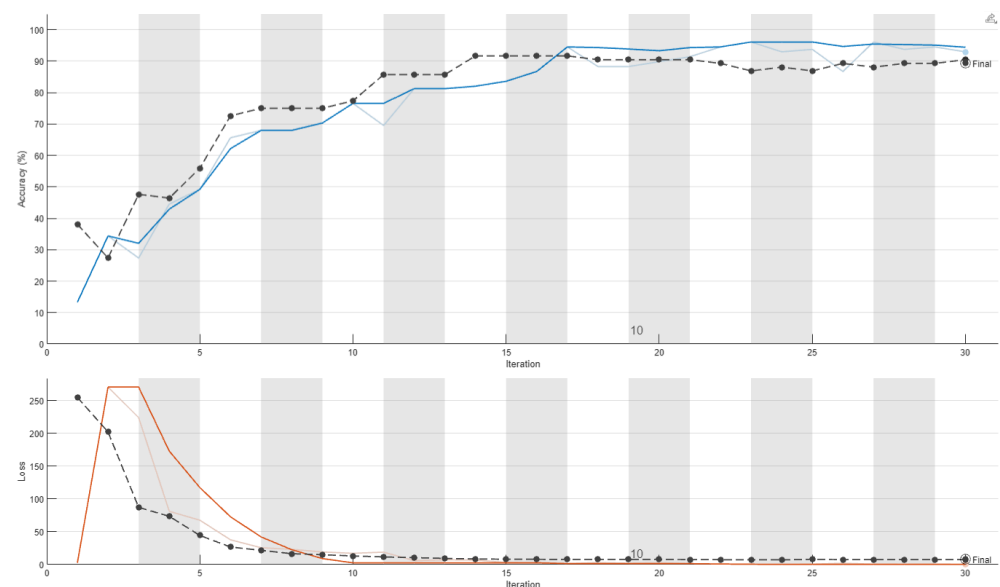
Figure 13. Training of CNN architecture proposed by Shahverdi et al. [81] without respect to humans on CSI-HAR [90] dataset. In the upper figure, the training accuracy is represented by the blue line, whereas test accuracy is depicted in black. In the lower figure, the training loss is indicated in red, and the test loss is shown in black.

Table 8. Comparison of results on CSI-HAR [90].

	Reported in [81]	Retrained w/o.r.t. Humans	Retrained w.r.t. Humans
Architecture	Accuracy	Accuracy	Accuracy
Plain RGB images	0.912	0.901	0.586
Canny	0.979	0.970	0.614
Sobel	0.971	0.962	0.604
Prewitt	0.961	0.954	0.597
LoG	0.975	0.966	0.610
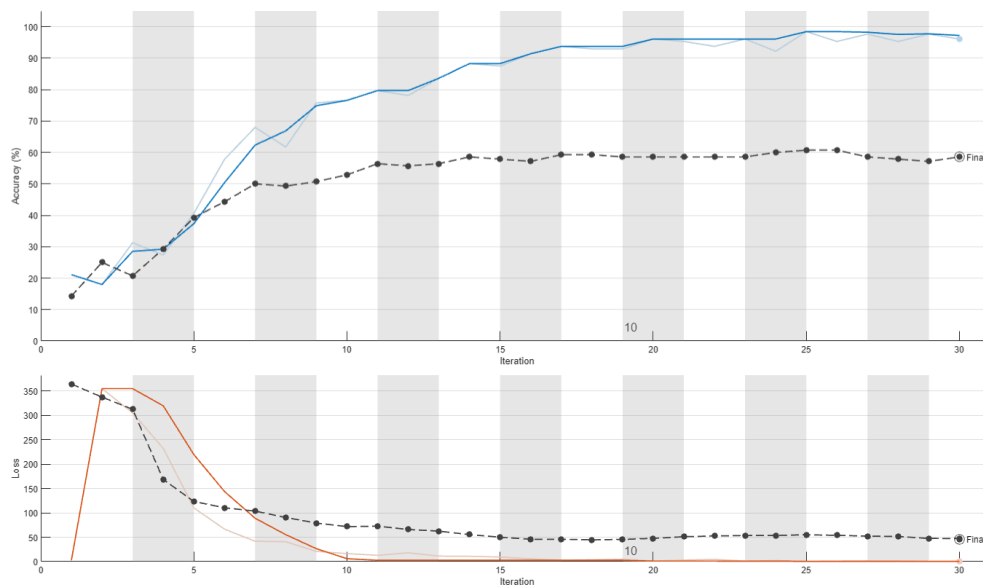
Figure 14. Training of CNN architecture proposed by Shahverdi et al. [81] with respect to humans on CSI-HAR [90] dataset. In the upper figure, the training accuracy is represented by the blue line, whereas the test accuracy is depicted in black. In the lower figure, the training loss is indicated in red, and the test loss is shown in black.

6. Discussion

The findings of this study underscore the critical importance of proper dataset partitioning in the development and evaluation of Wi-Fi CSI-based human action recognition systems. Through our analysis of three published methods, we have demonstrated that data leakage can significantly inflate performance metrics when the dataset is not partitioned with respect to individual subjects. This is particularly problematic, as it gives a false impression of a model's generalizability and robustness.

The results obtained from retraining without respect to individual subjects were slightly lower than those reported in the original studies. This discrepancy may be attributed to potential techniques such as data augmentation, K-fold cross-validation, or regularization that were not disclosed in the original papers. Despite this minor difference, our analysis clearly demonstrated the presence of data leakage, which severely undermines the reliability of the reported results.

Retraining with subject-based partitioning yielded results that were approximately two-thirds of the original reported outcomes in the most favorable case. While these results may appear discouraging, they reflect a more accurate assessment of the models' ability to generalize to new, unseen subjects. This highlights the necessity of employing correct data splitting methods, even if it results in lower performance metrics, as it ensures the validity and integrity of the evaluation process.

The insights gained from this study have several implications for future research. First, researchers must adopt rigorous dataset partitioning strategies to prevent data leakage and ensure the development of reliable and generalizable models. Transparency in reporting

experimental setups, including training parameters and evaluation protocols, is essential to enable the reproducibility and accurate comparison of results [92]. By addressing these issues, we aim to contribute to the establishment of best practices in the field of Wi-Fi CSI-based human action recognition, ultimately leading to the development of more robust and trustworthy systems.

7. Conclusions

This study highlights the critical issue of data leakage in Wi-Fi CSI-based HAR systems. By analyzing three published methods, we demonstrated how improper dataset partitioning, specifically, the failure to partition data with respect to individual subjects, can lead to significantly inflated performance metrics. Our findings reveal that models evaluated with data leakage present an unrealistically high accuracy, which is not reflective of their true generalizability. Upon re-evaluating the methods with correct subject-based partitioning, we observed a notable decline in performance, underscoring the importance of proper dataset management. These results, approximately two-thirds of the originally reported outcomes, provide a more accurate assessment of the models' capabilities in real-world scenarios. Additionally, our study found that the impact of various preprocessing techniques, such as Canny, Sobel, Prewitt, and LoG filtering, is less significant when appropriate data splitting is applied. As a consequence, there is still a lot of space for future research despite the high performance metrics found in the literature. Further, we pointed out that several activities seem to be very challenging for deep architectures to correctly recognize based on Wi-Fi CSI signals. In summary, the field still holds challenges that should be addressed by researchers and engineers.

Our work builds on previous research, where we identified data leakage in another method [11], further reinforcing the need for standardized evaluation practices in this field. We advocate for greater transparency in reporting experimental setups and the adoption of best practices in dataset partitioning to ensure the development of robust and generalizable models. By addressing these methodological flaws, we aim to pave the way for more reliable Wi-Fi CSI-based HAR systems. Future research should continue to focus on eliminating data leakage and improving the reproducibility and comparability of results across studies. Through these efforts, we can advance the field towards more trustworthy and effective applications in various real-world contexts.

Abbreviations

The following abbreviations are used in this manuscript:

CNN	convolutional neural network
CSI	channel state information
GADF	Gramian angular difference field
GAF	Gramian angular field

GASF	Gramian angulat summation field
GPU	graphics processing unit
HAR	human action recognition
IEEE	Institute of Electrical and Electronics Engineers
LDA	linear discriminant analysis
LSTM	long short-term memory
MIMO	multiple-input multiple-output
PCA	principal component analysis
ReLU	rectified linear unit
RSSI	received signal strength indicator
SGDM	stochastic gradient descent with momentum
SVM	support vector machine

References

1. Yadav, S.K.; Sai, S.; Gundewar, A.; Rathore, H.; Tiwari, K.; Pandey, H.M.; Mathur, M. CSITime: Privacy-preserving human activity recognition using WiFi channel state information. *Neural Netw.* **2022**, *146*, 11–21. [CrossRef] [PubMed]
2. Zhang, H.B.; Zhang, Y.X.; Zhong, B.; Lei, Q.; Yang, L.; Du, J.X.; Chen, D.S. A comprehensive survey of vision-based human action recognition methods. *Sensors* **2019**, *19*, 1005. [CrossRef] [PubMed]
3. Yan, H.; Zhang, Y.; Wang, Y.; Xu, K. WiAct: A passive WiFi-based human activity recognition system. *IEEE Sens. J.* **2019**, *20*, 296–305. [CrossRef]
4. Wang, Z.; Jiang, K.; Hou, Y.; Dou, W.; Zhang, C.; Huang, Z.; Guo, Y. A survey on human behavior recognition using channel state information. *IEEE Access* **2019**, *7*, 155986–156024. [CrossRef]
5. Cheng, X.; Huang, B.; Zong, J. Device-free human activity recognition based on GMM-HMM using channel state information. *IEEE Access* **2021**, *9*, 76592–76601. [CrossRef]
6. Yousefi, S.; Narui, H.; Dayal, S.; Ermon, S.; Valaee, S. A survey on behavior recognition using WiFi channel state information. *IEEE Commun. Mag.* **2017**, *55*, 98–104. [CrossRef]
7. Khan, U.M.; Kabir, Z.; Hassan, S.A. Wireless health monitoring using passive WiFi sensing. In Proceedings of the 2017 13th International Wireless Communications and Mobile Computing Conference (IWCMC), Valencia, Spain, 18–23 June 2017; pp. 1771–1776.
8. Jiang, H.; Cai, C.; Ma, X.; Yang, Y.; Liu, J. Smart home based on WiFi sensing: A survey. *IEEE Access* **2018**, *6*, 13317–13325. [CrossRef]
9. Sruthy, S.; George, S.N. WiFi enabled home security surveillance system using Raspberry Pi and IoT module. In Proceedings of the 2017 IEEE International Conference on Signal Processing, Informatics, Communication and Energy Systems (SPICES), Kollam, Kerala, India, 8–10 August 2017; pp. 1–6.
10. Rallapalli, S.; Ganesan, A.; Chintalapudi, K.; Padmanabhan, V.N.; Qiu, L. Enabling physical analytics in retail stores using smart glasses. In Proceedings of the 20th Annual International Conference on Mobile Computing and Networking, Maui, HI, USA, 7–11 September 2014; pp. 115–126.
11. Varga, D. Critical Analysis of Data Leakage in WiFi CSI-Based Human Action Recognition Using CNNs. *Sensors* **2024**, *24*, 3159. [CrossRef] [PubMed]
12. Sun, Z.; Ke, Q.; Rahmani, H.; Bennamoun, M.; Wang, G.; Liu, J. Human action recognition from various data modalities: A review. *IEEE Trans. Pattern Anal. Mach. Intell.* **2022**, *45*, 3200–3225. [CrossRef]
13. Wu, X.; Chu, Z.; Yang, P.; Xiang, C.; Zheng, X.; Huang, W. TW-See: Human activity recognition through the wall with commodity Wi-Fi devices. *IEEE Trans. Veh. Technol.* **2018**, *68*, 306–319. [CrossRef]
14. Wang, Y.; Liu, J.; Chen, Y.; Gruteser, M.; Yang, J.; Liu, H. E-eyes: Device-free location-oriented activity identification using fine-grained wifi signatures. In Proceedings of the 20th Annual International Conference on Mobile Computing and Networking, Maui, HI, USA, 7–11 September 2014; pp. 617–628.
15. Wang, W.; Liu, A.X.; Shahzad, M.; Ling, K.; Lu, S. Device-free human activity recognition using commercial WiFi devices. *IEEE J. Sel. Areas Commun.* **2017**, *35*, 1118–1131. [CrossRef]
16. Guo, L.; Wang, L.; Liu, J.; Zhou, W.; Lu, B. HuAc: Human activity recognition using crowdsourced WiFi signals and skeleton data. *Wirel. Commun. Mob. Comput.* **2018**, *2018*, 1–15. [CrossRef]
17. Zhang, Z. Microsoft kinect sensor and its effect. *IEEE Multimed.* **2012**, *19*, 4–10. [CrossRef]
18. Wang, F.; Feng, J.; Zhao, Y.; Zhang, X.; Zhang, S.; Han, J. Joint activity recognition and indoor localization with WiFi fingerprints. *IEEE Access* **2019**, *7*, 80058–80068. [CrossRef]
19. Alazrai, R.; Awad, A.; Baha'A, A.; Hababeh, M.; Daoud, M.I. A dataset for Wi-Fi-based human-to-human interaction recognition. *Data Brief* **2020**, *31*, 105668. [CrossRef] [PubMed]
20. Brunton, S.L.; Kutz, J.N. *Data-Driven Science and Engineering: Machine Learning, Dynamical Systems, and Control*; Cambridge University Press: Cambridge, UK, 2022.
21. Wang, G.; Zou, Y.; Zhou, Z.; Wu, K.; Ni, L.M. We can hear you with Wi-Fi! In Proceedings of the 20th Annual International Conference on Mobile Computing and Networking, Maui, HI, USA, 7–11 September 2014; pp. 593–604.

22. Tan, S.; Yang, J. WiFinger: Leveraging commodity WiFi for fine-grained finger gesture recognition. In Proceedings of the 17th ACM International Symposium on Mobile ad hoc Networking and Computing, Athens, Greece, 14–17 October 2016; pp. 201–210.

23. Pu, Q.; Gupta, S.; Gollakota, S.; Patel, S. Whole-home gesture recognition using wireless signals. In Proceedings of the 19th Annual International Conference on Mobile Computing & Networking, Miami, FL, USA, 30 September–4 October 2013; pp. 27–38.

24. Adib, F.; Kabelac, Z.; Katabi, D.; Miller, R.C. 3D tracking via body radio reflections. In Proceedings of the 11th USENIX Symposium on Networked Systems Design and Implementation (NSDI 14), Seattle, WA, USA, 2–4 April 2014; pp. 317–329.

25. Adib, F.; Katabi, D. See through walls with WiFi! In Proceedings of the ACM SIGCOMM 2013 Conference on SIGCOMM, Hong Kong, China, 12–16 August 2013; pp. 75–86.

26. Liu, X.; Cao, J.; Tang, S.; Wen, J. Wi-sleep: Contactless sleep monitoring via wifi signals. In Proceedings of the 2014 IEEE Real-Time Systems Symposium, Rome, Italy, 2–5 December 2014; pp. 346–355.

27. Chen, C.; Shu, Y.; Shu, K.I.; Zhang, H. WiTT: Modeling and the evaluation of table tennis actions based on WIFI signals. In Proceedings of the 2018 24th International Conference on Pattern Recognition (ICPR), Beijing, China, 20–24 August 2018; pp. 3100–3107.

28. Wang, Y.; Wu, K.; Ni, L.M. Wifall: Device-free fall detection by wireless networks. *IEEE Trans. Mob. Comput.* **2016**, *16*, 581–594. [CrossRef]

29. Chang, C.C.; Lin, C.J. LIBSVM: A library for support vector machines. *ACM Trans. Intell. Syst. Technol. (TIST)* **2011**, *2*, 1–27. [CrossRef]

30. Wang, H.; Zhang, D.; Wang, Y.; Ma, J.; Wang, Y.; Li, S. RT-Fall: A real-time and contactless fall detection system with commodity WiFi devices. *IEEE Trans. Mob. Comput.* **2016**, *16*, 511–526. [CrossRef]

31. Hu, Y.; Zhang, F.; Wu, C.; Wang, B.; Liu, K.R. DeFall: Environment-independent passive fall detection using WiFi. *IEEE Internet Things J.* **2021**, *9*, 8515–8530. [CrossRef]

32. Duan, P.; Li, J.; Jiao, C.; Cao, Y.; Kong, J. WiBFall: A Device-Free Fall Detection Model for Bathroom. In Proceedings of the International Conference on Mobile Networks and Management, Chiba, Japan, 27–29 October 2021; pp. 182–193.

33. Chen, S.; Yang, W.; Xu, Y.; Geng, Y.; Xin, B.; Huang, L. AFall: Wi-Fi-based device-free fall detection system using spatial angle of arrival. *IEEE Trans. Mob. Comput.* **2022**, *22*, 4471–4484. [CrossRef]

34. Zhou, Q.; Xing, J.; Li, J.; Yang, Q. A device-free number gesture recognition approach based on deep learning. In Proceedings of the 2016 12th International Conference on Computational Intelligence and Security (CIS), Seville, Spain, 13–15 May 2016; pp. 57–63.

35. Wang, X.; Gao, L.; Mao, S.; Pandey, S. DeepFi: Deep learning for indoor fingerprinting using channel state information. In Proceedings of the 2015 IEEE Wireless Communications and Networking Conference (WCNC), New Orleans, LA, USA, 9–12 March 2015; pp. 1666–1671.

36. Wang, X.; Gao, L.; Mao, S. PhaseFi: Phase fingerprinting for indoor localization with a deep learning approach. In Proceedings of the 2015 IEEE Global Communications Conference (GLOBECOM), San Diego, CA, USA, 6–10 December 2015; pp. 1–6.

37. Wang, F.; Gong, W.; Liu, J. On spatial diversity in WiFi-based human activity recognition: A deep learning-based approach. *IEEE Internet Things J.* **2018**, *6*, 2035–2047. [CrossRef]

38. Ronneberger, O.; Fischer, P.; Brox, T. U-net: Convolutional networks for biomedical image segmentation. In *Medical Image Computing and Computer-Assisted Intervention, Proceedings of the MICCAI 2015: 18th International Conference, Munich, Germany, 5–9 October 2015*; Proceedings, Part III 18; Springer: Berlin/Heidelberg, Germany, 2015; pp. 234–241.

39. Wang, F.; Song, Y.; Zhang, J.; Han, J.; Huang, D. Temporal unet: Sample level human action recognition using wifi. *arXiv* **2019**, arXiv:1904.11953.

40. Chen, Z.; Zhang, L.; Jiang, C.; Cao, Z.; Cui, W. WiFi CSI based passive human activity recognition using attention based BLSTM. *IEEE Trans. Mob. Comput.* **2018**, *18*, 2714–2724. [CrossRef]

41. Huang, S.; Wang, D.; Zhao, R.; Zhang, Q. Wiga: A wifi-based contactless activity sequence recognition system based on deep learning. In Proceedings of the 2019 15th International Conference on Mobile Ad-Hoc and Sensor Networks (MSN), Shenzhen, China, 11–13 December 2019; pp. 69–74.

42. Sheng, B.; Xiao, F.; Sha, L.; Sun, L. Deep spatial–temporal model based cross-scene action recognition using commodity WiFi. *IEEE Internet Things J.* **2020**, *7*, 3592–3601. [CrossRef]

43. Sherstinsky, A. Fundamentals of recurrent neural network (RNN) and long short-term memory (LSTM) network. *Phys. D Nonlinear Phenom.* **2020**, *404*, 132306. [CrossRef]

44. Denil, M.; Bazzani, L.; Larochelle, H.; de Freitas, N. Learning where to attend with deep architectures for image tracking. *Neural Comput.* **2012**, *24*, 2151–2184. [CrossRef]

45. Graves, A.; Fernández, S.; Schmidhuber, J. Bidirectional LSTM networks for improved phoneme classification and recognition. In Proceedings of the International Conference on Artificial Neural Networks, Warsaw, Poland, 11–15 September 2005; pp. 799–804.

46. Zou, H.; Yang, J.; Prasanna Das, H.; Liu, H.; Zhou, Y.; Spanos, C.J. WiFi and vision multimodal learning for accurate and robust device-free human activity recognition. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops, Long Beach, CA, USA, 16–20 June 2019.

47. Ye, H.; Wu, Z.; Zhao, R.W.; Wang, X.; Jiang, Y.G.; Xue, X. Evaluating two-stream CNN for video classification. In Proceedings of the 5th ACM on International Conference on Multimedia Retrieval, Shanghai, China, 23–26 June 2015; pp. 435–442.

48. Memmesheimer, R.; Theisen, N.; Paulus, D. Gimme signals: Discriminative signal encoding for multimodal activity recognition. In Proceedings of the 2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Las Vegas, NV, USA, 25–29 October 2020; pp. 10394–10401.

49. Yue, R.; Tian, Z.; Du, S. Action recognition based on RGB and skeleton data sets: A survey. *Neurocomputing* **2022**, *512*, 287–306. [CrossRef]

50. Zou, Q.; Ni, L.; Wang, Q.; Li, Q.; Wang, S. Robust gait recognition by integrating inertial and RGBD sensors. *IEEE Trans. Cybern.* **2017**, *48*, 1136–1150. [CrossRef] [PubMed]

51. Vrigkas, M.; Nikou, C.; Kakadiaris, I.A. A review of human activity recognition methods. *Front. Robot. AI* **2015**, *2*, 28. [CrossRef]

52. Pareek, P.; Thakkar, A. A survey on video-based human action recognition: Recent updates, datasets, challenges, and applications. *Artif. Intell. Rev.* **2021**, *54*, 2259–2322. [CrossRef]

53. Kong, Y.; Fu, Y. Human action recognition and prediction: A survey. *Int. J. Comput. Vis.* **2022**, *130*, 1366–1401. [CrossRef]

54. Chen, K.; Zhang, D.; Yao, L.; Guo, B.; Yu, Z.; Liu, Y. Deep learning for sensor-based human activity recognition: Overview, challenges, and opportunities. *ACM Comput. Surv. (CSUR)* **2021**, *54*, 1–40. [CrossRef]

55. Liu, J.; Liu, H.; Chen, Y.; Wang, Y.; Wang, C. Wireless sensing for human activity: A survey. *IEEE Commun. Surv. Tutorials* **2019**, *22*, 1629–1645. [CrossRef]

56. Hannun, A.; Guo, C.; van der Maaten, L. Measuring data leakage in machine-learning models with fisher information. In Proceedings of the Uncertainty in Artificial Intelligence, Online, 27–29 July 2021; pp. 760–770.

57. Stock, A.; Gregr, E.J.; Chan, K.M. Data leakage jeopardizes ecological applications of machine learning. *Nat. Ecol. Evol.* **2023**, *7*, 1743–1745. [CrossRef]

58. Yang, M.; Zhu, J.J.; McGaughey, A.; Zheng, S.; Priestley, R.D.; Ren, Z.J. Predicting extraction selectivity of acetic acid in pervaporation by machine learning models with data leakage management. *Environ. Sci. Technol.* **2023**, *57*, 5934–5946. [CrossRef] [PubMed]

59. Samala, R.K.; Chan, H.P.; Hadjiiski, L.; Koneru, S. Hazards of data leakage in machine learning: A study on classification of breast cancer using deep neural networks. In Proceedings of the Medical Imaging 2020: Computer-Aided Diagnosis, Houston, TX, USA, 16–19 February 2020; Volume 11314, pp. 279–284.

60. Rosenblatt, M.; Tejavibulya, L.; Jiang, R.; Noble, S.; Scheinost, D. The effects of data leakage on connectome-based machine learning models. *bioRxiv* **2023**, bioRxiv:2023.06.09.544383.

61. Rosenblatt, M.; Tejavibulya, L.; Jiang, R.; Noble, S.; Scheinost, D. Data leakage inflates prediction performance in connectome-based machine learning models. *Nat. Commun.* **2024**, *15*, 1829. [CrossRef] [PubMed]

62. Dong, Q. Leakage prediction in machine learning models when using data from sports wearable sensors. *Comput. Intell. Neurosci.* **2022**, *2022*, 5314671. [CrossRef]

63. Moghaddam, A.K.; Zincir-Heywood, N. Exploring data leakage in encrypted payload using supervised machine learning. In Proceedings of the 15th International Conference on Availability, Reliability and Security, virtual event, 25–28 August 2020; pp. 1–10.

64. Poldrack, R.A.; Huckins, G.; Varoquaux, G. Establishment of best practices for evidence for prediction: A review. *JAMA Psychiatry* **2020**, *77*, 534–540. [CrossRef] [PubMed]

65. Kapoor, S.; Narayanan, A. Leakage and the reproducibility crisis in ML-based science. *arXiv* **2022**, arXiv:2207.07048.

66. Kapoor, S.; Narayanan, A. Leakage and the reproducibility crisis in machine-learning-based science. *Patterns* **2023**, *4*. [CrossRef]

67. Jiao, W.; Zhang, C. An Efficient Human Activity Recognition System Using WiFi Channel State Information. *IEEE Syst. J.* **2023**, *17*, 6687–6690. [CrossRef]

68. Xu, Z.; Lin, H. Quantum-Enhanced Forecasting: Leveraging Quantum Gramian Angular Field and CNNs for Stock Return Predictions. *arXiv* **2023**, arXiv:2310.07427.

69. Wang, Z.; Oates, T. Imaging time-series to improve classification and imputation. *arXiv* **2015**, arXiv:1506.00327.

70. Krizhevsky, A.; Sutskever, I.; Hinton, G.E. ImageNet classification with deep convolutional neural networks. *Commun. ACM* **2017**, *60*, 84–90. [CrossRef]

71. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep residual learning for image recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016; pp. 770–778.

72. Simonyan, K.; Zisserman, A. Very deep convolutional networks for large-scale image recognition. *arXiv* **2014**, arXiv:1409.1556.

73. Zhang, X.; Zhou, X.; Lin, M.; Sun, J. Shufflenet: An extremely efficient convolutional neural network for mobile devices. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–22 June 2018; pp. 6848–6856.

74. Kingma, D.P.; Ba, J. Adam: A method for stochastic optimization. *arXiv* **2014**, arXiv:1412.6980.

75. Jawad, S.K.; Alaziz, M. Human Activity and Gesture Recognition Based on WiFi Using Deep Convolutional Neural Networks. *Iraqi J. Electr. Electron. Eng.* **2022**, *18*, 110–116. [CrossRef]

76. Iandola, F.N.; Han, S.; Moskewicz, M.W.; Ashraf, K.; Dally, W.J.; Keutzer, K. SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and <0.5 MB model size. *arXiv* **2016**, arXiv:1602.07360.

77. Zhang, C.; Jiao, W. Imgfi: A high accuracy and lightweight human activity recognition framework using csi image. *IEEE Sens. J.* **2023**, *23*, 21966–21977. [CrossRef]

78. Casdagli, M. Recurrence plots revisited. *Phys. D Nonlinear Phenom.* **1997**, *108*, 12–44. [CrossRef]

79. Solomon, C.; Breckon, T. *Fundamentals of Digital Image Processing: A Practical Approach with Examples in Matlab*; John Wiley & Sons: Hoboken, NJ, USA, 2011.

80. Deng, J.; Dong, W.; Socher, R.; Li, L.J.; Li, K.; Fei-Fei, L. Imagenet: A large-scale hierarchical image database. In Proceedings of the 2009 IEEE Conference on Computer Vision and Pattern Recognition, Miami, FL, USA, 20–25 June 2009; pp. 248–255.

81. Shahverdi, H.; Nabati, M.; Fard Moshiri, P.; Asvadi, R.; Ghorashi, S.A. Enhancing CSI-based human activity recognition by edge detection techniques. *Information* **2023**, *14*, 404. [CrossRef]

82. Greenacre, M.; Groenen, P.J.; Hastie, T.; d'Enza, A.I.; Markos, A.; Tuzhilina, E. Principal component analysis. *Nat. Rev. Methods Prim.* **2022**, *2*, 100. [CrossRef]

83. Balakrishnama, S.; Ganapathiraju, A. Linear discriminant analysis-a brief tutorial. *Inst. Signal Inf. Process.* **1998**, *18*, 1–8.

84. Srivastava, N.; Hinton, G.; Krizhevsky, A.; Sutskever, I.; Salakhutdinov, R. Dropout: A simple way to prevent neural networks from overfitting. *J. Mach. Learn. Res.* **2014**, *15*, 1929–1958.

85. Dubey, A.K.; Jain, V. Comparative study of convolution neural network's relu and leaky-relu activation functions. In *Applications of Computing, Automation and Wireless Systems in Electrical Engineering: Proceedings of MARC 2018*; Springer: Singapore, 2019; pp. 873–880.

86. Ketkar, N.; Moolayil, J.; Ketkar, N.; Moolayil, J. Introduction to pytorch. In *Deep Learning with Python: Learn Best Practices of Deep Learning Models with PyTorch*; Apress: Berkeley, CA, USA, 2021; pp. 27–91.

87. Oakden, T.; Kavakli, M. Performance Analysis of RTX Architecture in Virtual Production and Graphics Processing. In Proceedings of the 2022 IEEE 42nd International Conference on Distributed Computing Systems Workshops (ICDCSW), Bologna, Italy, 10–13 July 2022; pp. 215–220.

88. Guo, L.; Wang, L.; Lin, C.; Liu, J.; Lu, B.; Fang, J.; Liu, Z.; Shan, Z.; Yang, J.; Guo, S. Wiar: A public dataset for wifi-based activity recognition. *IEEE Access* **2019**, *7*, 154935–154945. [CrossRef]

89. Zhang, Y.; Zheng, Y.; Qian, K.; Zhang, G.; Liu, Y.; Wu, C.; Yang, Z. Widar3.0: Zero-effort cross-domain gesture recognition with wi-fi. *IEEE Trans. Pattern Anal. Mach. Intell.* **2021**, *44*, 8671–8688. [CrossRef]

90. Moshiri, P.F.; Shahbazian, R.; Nabati, M.; Ghorashi, S.A. A CSI-based human activity recognition using deep learning. *Sensors* **2021**, *21*, 7225. [CrossRef] [PubMed]

91. Gringoli, F.; Schulz, M.; Link, J.; Hollick, M. Free your CSI: A channel state information extraction platform for modern Wi-Fi chipsets. In Proceedings of the 13th International Workshop on Wireless Network Testbeds, Experimental Evaluation & Characterization, Los Cabos, Mexico, 25 October 2019; pp. 21–28.

92. Saupe, D.; Hahn, F.; Hosu, V.; Zingman, I.; Rana, M.; Li, S. Crowd workers proven useful: A comparative study of subjective video quality assessment. In Proceedings of the QoMEX 2016: 8th International Conference on Quality of Multimedia Experience, Lisbon, Portugal, 6–8 June 2016.