

Article

Data Valuation Algorithm for Inertial Measurement Unit-Based Human Activity Recognition

Yeon-Wook Kim ¹ and Sangmin Lee ^{1,2,*}¹ Department of Electrical and Computer Engineering, Inha University, Incheon 22212, Republic of Korea² Department of Smart Engineering Program in Biomedical Science & Engineering, Inha University, Incheon 22212, Republic of Korea* Correspondence: sanglee@inha.ac.kr; Tel.: +82-32-860-7420

Abstract: This paper proposes a data valuation algorithm for inertial measurement unit-based human activity recognition (IMU-based HAR) data based on meta reinforcement learning. Unlike previous studies that received feature-level input, the algorithm in this study added a feature extraction structure to the data valuation algorithm, and it can receive raw-level inputs and achieve excellent performance. As IMU-based HAR data are multivariate time-series data, the proposed algorithm incorporates an architecture capable of extracting both local and global features by inserting a transformer encoder after the one-dimensional convolutional neural network (1D-CNN) backbone in the data value estimator. In addition, the 1D-CNN-based stacking ensemble structure, which exhibits excellent efficiency and performance on IMU-based HAR data, is used as a predictor to supervise model training. The Berg balance scale (BBS) IMU-based HAR dataset and the public datasets, UCI-HAR, WISDM, and PAMAP2, are used for performance evaluation in this study. The valuation performance of the proposed algorithm is observed to be excellent on IMU-based HAR data. The rate of discovering corrupted data is higher than 96% on all datasets. In addition, classification performance is confirmed to be improved by the suppression of discovery of low-value data.

Keywords: data valuation algorithm; meta-reinforcement learning; deep learning; transformer; convolutional neural network; human activity recognition; inertial measurement unit



Citation: Kim, Y.-W.; Lee, S. Data Valuation Algorithm for Inertial Measurement Unit-Based Human Activity Recognition. *Sensors* **2023**, *23*, 184. <https://doi.org/10.3390/s23010184>

Academic Editors: ByoungChul Ko and Jong-Ha Lee

Received: 15 November 2022

Revised: 19 December 2022

Accepted: 20 December 2022

Published: 24 December 2022



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Deep learning algorithms have been used in various fields [1–3]. In the field of computer vision, deep learning algorithms are used to perform multi-modal learning to obtain useful information from images and texts, images and speech, and images and sensor signals [1]. In the field of audio analysis, deep learning algorithms are used for automatic speech recognition, audio enhancement, and audio generation [2]. In the field of natural language processing, deep learning algorithms are used to perform sentiment analysis and machine translation [3]. The training of these deep learning algorithms tasks requires large amounts of data. Using sufficient data during training prevents overfitting and enhances generalizability. Therefore, deep learning algorithms are widely used for big data analysis [4–6]. Unlike machine learning models that require handcrafted feature engineering, deep learning models can extract features and rules from data and output the desired signals or labels [7–9]. Therefore, both model structure and quality of training data play important roles in improving the performance of deep learning-based models.

The quality of training data is often degraded for various reasons. In the case of sensor data, artifacts of the sensor device or noise caused by the environment can degrade signal quality [10,11]. Moreover, manual labeling of collected data may be erroneous owing to mistakes or insufficient information [12,13]. Data collected via crawling may be unintentionally collected or incorrectly labeled [14,15]. Data labeled via crowdsourcing can also include labeling errors owing to human subjectivity or mistakes [16]. In turn, low-quality

data of the aforementioned types degrade the model performance [17]. The challenge lies in the fact that obtaining high-quality data is time-consuming and expensive [14].

Several studies have implemented robust models, even with low-quality data, to overcome this limitation [18–20]. Moreover, studies have been conducted to improve the performance of algorithms by improving the data quality. Liu [21] vectorized restaurant names and user comments in social networks and improved the low-quality data and data without location labels based on cosine similarity. The performance of the labeling model was improved using game theory [22]. Ju [23] proposed an algorithm for reducing the label noise of labeled medical images based on the Monte Carlo estimation method [24] and a CNN model.

Studies have also been conducted on data valuation algorithms to improve the quality of the training data. Data valuation algorithms evaluate the value of each training sample to be used as its training weight. Leave-one-out (LOO) [25] is the most commonly used data valuation method—it evaluates each sample by adopting the difference between the performance of the model, including and excluding the sample as the sample value. As the computational complexity of LOO increases linearly concerning the number of training samples, it is not suitable for large datasets [25]. Data Shapley [26] is another data valuation algorithm inspired by game theory [22]. It uses marginal performance improvement as the data value after calculating the performance on all possible subsets of the training data. Its computational complexity increases exponentially with the number of training data as it requires training on all possible subsets. Monte Carlo sampling [24] can be used for approximation to reduce the computational complexity. However, it exhibits high computational complexity itself, and approximation introduces certain limitations.

Recently, a meta-learning-based algorithm was proposed that addresses the aforementioned limitations of high computational complexity and approximation. Ren et al. [27] proposed a robust algorithm for low-quality data by adjusting the weight of the batch-size training data for each gradient step using the validation set. Hendrycks [28] corrected the labels of corrupted label data using a clean validation set and re-trained the model using the corrected training data. Saeed [29] used a neural network-based task predictor for image segmentation and classification to update the neural network-based image quality assessment (IQA) controller for medical image data. The authors performed meta-reinforcement learning for newly added data or meta-task data to fine-tune the IQA controller network using the task performance of the predictor. Yoon [30] proposed a deep learning-based data valuation algorithm using reinforcement learning by combining a predictor with a data value estimator (DVE). The DVE was trained using meta-reinforcement learning using the task performance of the predictor. This method exhibited better performance and efficiency than LOO [25] and Data Shapley [26]. Previous data valuation algorithm studies [26,29,30] primarily dealt with public image datasets and insufficiently dealt with other types of data, such as time series data and data with a small number of samples.

In this paper, we propose a data-valuation algorithm based on meta-reinforcement learning for inertial measurement unit-based human activity recognition (hereinafter, IMU-based HAR) data. The IMU comprises a three-axis acceleration sensor and a three-axis gyroscope sensor and measures the inertia applied to the unit based on the captured motion information. The IMU-based HAR algorithm, which is a type of pattern recognition algorithm, recognizes the type and quality of motion based on the IMU data collected using wearable devices. Previous studies on data valuation algorithms have primarily utilized vision-based public datasets [26,29,30]. However, IMU-based HAR data have not been investigated yet. Unlike [30], which required feature-level input, the proposed algorithm adds a feature extraction structure to the data valuation algorithm, enabling the utilization of raw-level inputs. In the algorithm proposed in [30], a pre-trained model was required to accept feature-level input data, which required training data of sufficient volume and quality. Thus, constructing good pre-trained models may be difficult in some cases. The proposed algorithm does not suffer from this limitation. Meanwhile, in previous studies, pre-trained models were used to generate feature-level data to train the

network for prediction. However, if a feature extraction structure is included in the data valuation algorithm, as in the proposed algorithm, its network is trained for the purpose of data valuation.

The proposed data valuation algorithm incorporates a structure suitable for IMU-based HAR data within it. It comprises a DVE that receives the training data input and a predictor that supervises the output of the DVE during model training. A feature extraction structure suitable for IMU-based HAR data is added to the DVE to enable it to accept raw-level inputs. Both local and global features can be extracted from the IMU-based HAR data using the feature extraction structure, where a transformer encoder is inserted after the 1D-CNN backbone. A stacking ensemble structure, including a double-head 1D-CNN, which exhibits good performance and efficiency on IMU-based HAR data, is used as the predictor. Four IMU-based HAR datasets are used for model evaluation—the Berg balance scale (BBS) HAR data collected at Inha University Hospital and the public IMU-based HAR datasets from the University of California, Irvine, human activity recognition using a smartphone dataset (UCI-HAR), wireless sensor data mining (WISDM), and the Physical Activity Monitoring dataset (PAMAP2). Corrupted data are generated by contaminating the labels of 20% of the training data to evaluate the algorithm. On all IMU-based HAR datasets, excellent performance is observed concerning the ratio of finding corrupted data in low-value data, exceeding 96%. In addition, a classification performance is observed to improve on all IMU-based HAR datasets when low-value data are removed from the training data. This indicates that the proposed algorithm evaluates the data adequately.

2. Materials and Methods

2.1. Proposed Algorithm

2.1.1. Structure of the Proposed Data Valuation Algorithm

The algorithm comprises a DVE and a predictor that supervises the output of the DVE during training. The structure of the algorithm is illustrated in Figure 1.

In the case of a training sample input of batch size = B , the DVE extracts appropriate features from it, concatenates it with label information, and refines the information of the input data using a multi-layer perceptron (MLP) structure comprising five dense layers. All dense layers are structurally identical, with 100 perceptrons and the ReLU activation function. The final dense layer in the MLP is concatenated with marginal information which is the degree of contamination of training data. The marginal information is given by $m(\mathbf{x}, y) = |y - f_v(\mathbf{x})|$, where f_v denotes a predictor pretrained using validation data. Subsequently, a dense layer is placed, and the selection probability, $h_\varphi(\mathbf{x}, y)$, of the corresponding data is output as softmax. The selection probability is equal to the value of the corresponding sample. Corresponding to training data, $D = \{(\mathbf{x}, y)\}_{i=1}^N \sim P$, the sampler uses the polynomial distribution $h_\varphi(D)$ obtained by the DVE to choose the selection vector $s = \{s_1, s_2, \dots, s_B\}$. The probability of outputting the selection vector, s , is $\pi_\varphi(D, s) = \prod_{i=1}^N [h_\varphi(\mathbf{x}_i, y_i)^{s_i} \cdot (1 - h_\varphi(\mathbf{x}_i, y_i))^{1-s_i}]$. The DVE output is passed as a training weight for each sample of the predictor model.

The predictor, f_θ , is trained to minimize a weighted loss function, L_f , on the training dataset, D . Equation (1) express this. Cross-entropy is used as the loss function, and θ denotes a parameter of the predictor model.

$$f_\theta = \operatorname{argmin}_{f \in F} \frac{1}{N} \sum_{i=1}^N h_\varphi(\mathbf{x}_i, y_i) \cdot \nabla L_f(\hat{f}_\theta(\mathbf{x}_i), y_i) \quad (1)$$

The task performance of trained predictor is used as the loss, L_h . At this time, to calculate the task performance, clean validation data $D^v = \{(\mathbf{x}^v, y^v)\}_{k=1}^L \sim P^t$ are used. The DVE loss is trained using a gradient-based method. The loss is transmitted to the DVE

as a reward. The loss for DVE training obtained by multiplying the probability and reward corresponding to the sample can be expressed as the following expected value:

$$\begin{aligned} \hat{l}(\varphi) &= E_{(x^v, y^v) \sim P^t, s \sim \pi_\varphi(D, \cdot)} [L_h(f_\theta(x^v), y^v)] \\ &= \int P^t(x^v) \sum_{s \in [0,1]^N} \pi_\varphi(D, s) \cdot [L_h(f_\theta(x^v), y^v)] dx^v \end{aligned} \tag{2}$$

At this time, the agent is DVE, the action is the data selection process, and the environment encompasses training and evaluation of the predictor.

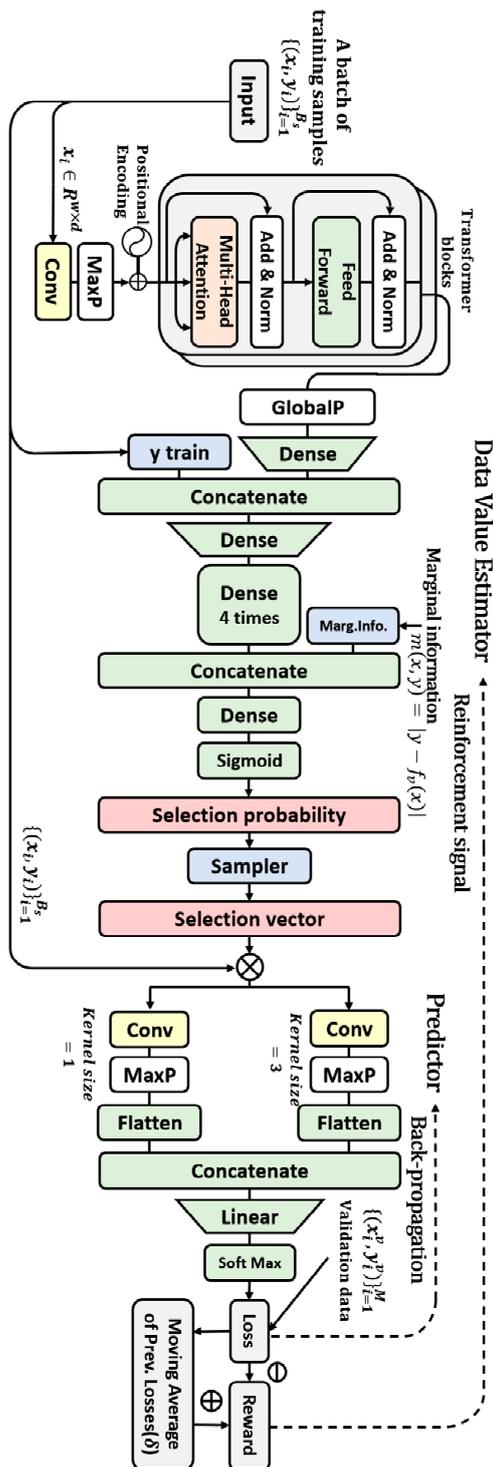


Figure 1. Structure of the proposed algorithm.

2.1.2. Structure for IMU-Based HAR in Proposed Data Valuation Algorithm

The proposed data valuation algorithm introduces appropriate deep learning structures into an existing algorithm [30] to achieve a good performance on IMU-based HAR data. A structure for extracting the features of IMU-based HAR data is inserted into the data input part of the DVE to enable raw-level, IMU-based HAR data to be accepted as input. The predictor uses a module with good efficiency and performance on IMU-based HAR data.

The feature extraction structure extracts both local and global features by inserting a transformer encoder after the 1D-CNN backbone. The filter size of the 1D-CNN layer is taken to be 64, its kernel size is taken to be 3, and the GeLU activation function is used. Two self-attention heads are used in the transformer block, 256 perceptrons are used in the feed-forward layer, a dropout of 0.1 is used, GeLU is used as the activation function, and two structurally identical transformers are placed in a row. In previous studies on IMU-based HAR data [31–33], models comprising a recurrent neural network (RNN) series after the 1D-CNN exhibited good performance. The feature extraction structure used in this study is inspired by those used in previous studies [31–33]. Rather than using an RNN-series model, a transformer encoder block is used to perform a similar role. The latter is superior to the former in terms of computational efficiency [34].

A 1D-CNN-based stacking ensemble structure model [35–38] that exhibits good performance and efficiency in an inertial sensor-based HAR algorithm is used as the predictor. Structurally, it consists of a simple dense layer classifier after a double-head 1D-CNN, and the kernel sizes of the two heads are taken to be 1 and 3 to extract different features. The filter size of the 1D CNN layer is taken to be 64, and ReLU is used as the activation function.

2.2. Evaluation Datasets

In this study, four IMU-based HAR datasets are used to evaluate the proposed algorithm—the BBS HAR data collected together with the Department of Rehabilitation Medicine at Inha University Hospital and public datasets, namely, UCI-HAR, WISDM, and PAMAP2 data.

2.2.1. BBS HAR Data

The BBS HAR dataset comprises IMU-based HAR data recorded by introducing a wearable inertial measurement unit (IMU) into a BBS, a balanced assessment. The BBS is the balance assessment designed to evaluate the balance ability of the elderly and is known to be highly reliable, even for patients with brain diseases [39,40]. In BBS, subjects are asked to perform 14 static and dynamic tasks, each of which is scored. The balance ability of the subject is evaluated on the basis of the total score [41].

The data were collected from the Department of Rehabilitation Medicine at Inha University Hospital. The experimental design was approved by the Institutional Review Board. In aggregate, 53 patients aged 50–80 years (male: 31, female: 22) with brain disease and three healthy individuals in their late 20s participated in the experiment. The healthy participants imitated the motions of the patients and performed all motions with scores between 0 and 4.

Noraxon's myoMotion, which is a multichannel wireless IMU system, is used for the experiment. This system is certified to be an ISO 13,485 compliant (Registration # MED-0037b) and an FDA 510 K compliant (Registration number #2098416) medical device. IMU modules are attached to the human body using Velcro bands. The IMU modules transmit data wirelessly to receivers, which are connected to a computer via a USB. The system uses a type of PC software for recording and management hardware. If a webcam is connected to the PC, video data synchronized with IMU data can be recorded, which can be used to label IMU motion data. Figure 2 illustrates the software and equipment of the Noraxon's myoMotion.

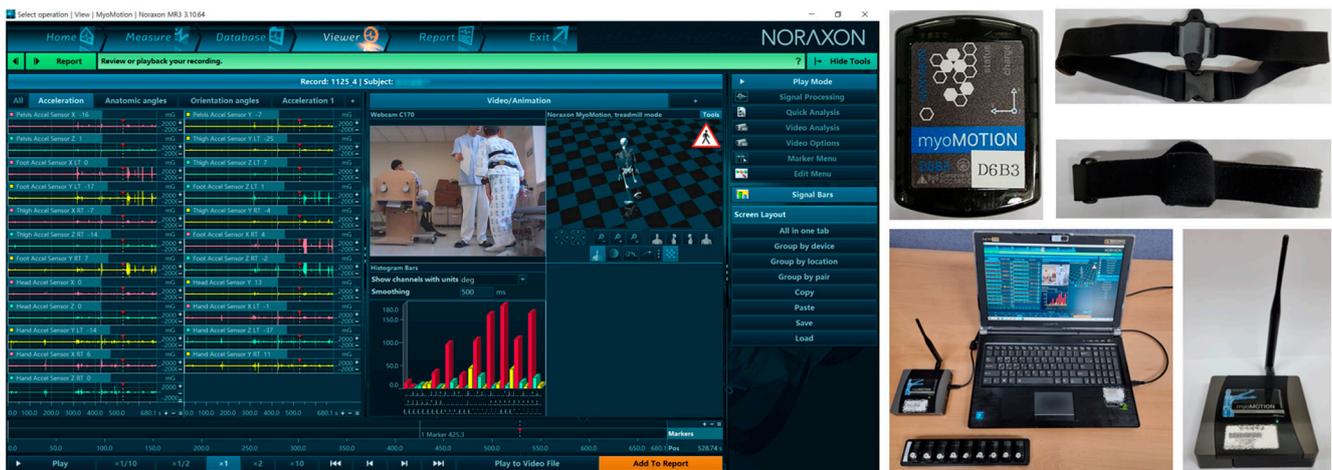


Figure 2. Software and equipment of the multichannel wireless IMU system.

Eight IMUs are used in the experiment and worn on the forehead, left and right wrists, left and right upper hips, back of the left and right ankles, and back. Each IMU recorded the values of accelerometer X, Y, Z, roll, pitch, yaw, rotation X, Y, Z at a sampling rate of 100 Hz. The rotation refers to the number of rotations and is calculated by cumulatively aggregating the rotation angles of the sensor. Figure 3 depicts a photograph of the BBS experimental environment and the IMU attachment locations.

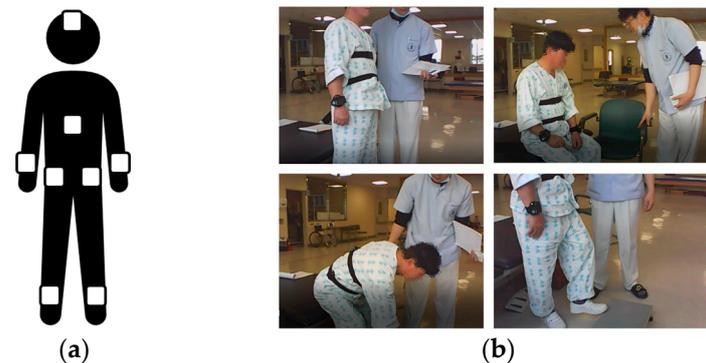


Figure 3. IMU attachment locations and photograph of experiment: (a) IMU attachment locations; (b) photograph of the BBS experiment.

In [35], a deep-learning-based BBS score recognition algorithm was proposed. We adopt the data pre-processing methodology of the aforementioned algorithm, comprising data augmentation based on the oversampling technique, data downsampling, normalization, and zero-padding [35].

2.2.2. UCI-HAR

The UCI-HAR dataset comprises IMU-based HAR data obtained using inertial sensors embedded in smartphones and were devised by Anguita [42]. The participants performed six motions in aggregate—“walking”, “walking upstairs”, “walking downstairs”, “sitting”, “standing”, and “lying down”—while wearing a smartphone on their waist. A total of 30 participants aged 19–48 years participated in the experiment. Motion data were recorded using a 3-axis gyroscope and 3-axis accelerometer at a sampling rate of 50 Hz. A sliding window was applied to the data for real-time recognition. The window size was taken to be 128, with an overlap of 50%. Data augmentation was performed using the same method as that in a previous study [37].

2.2.3. WISDM

The WISDM dataset comprises HAR data obtained using inertial sensors embedded in smartphones and were devised by Kwapisz [43]. Each participant performed six movements—“walking”, “jogging”, “ascending stairs”, “descending stairs”, “sitting”, and “standing”—with a smartphone in the front pocket of their trousers. A total of 36 people participated in the experiment, and the 3-axis accelerometer data were recorded at a sampling rate of 20 Hz. For real-time recognition, a sliding window was applied to WISDM data with a window size of 80 and an overlap of 50%. Data augmentation was performed using the same method as in a previous study [37].

2.2.4. PAMAP2

The PAMAP2 dataset comprises IMU-based HAR data collected by Reiss [44] from test participants using three wearable IMUs on their hands, chest, and ankles and a heart rate sensor. The test participants performed 12 movements commonly performed in daily life—“lying”, “sitting”, “standing”, “walking”, “running”, “cycling”, “Nordic walking”, “ascending stairs”, “descending stairs”, “vacuum cleaning”, “ironing”, and “rope jumping”—and six optional movements—“watching TV”, “computer work”, “car driving”, “folding laundry”, “house cleaning”, and “playing soccer.” In this study, 12 types of data corresponding to actions undertaken in daily life are used. Nine participants aged 27–32 years participated in the test. The 3-axis gyroscope, 3-axis accelerometer, and 3-axis geomagnetic and temperature sensor data were recorded at a sampling rate of 100 Hz, and the heart rate data were recorded at a sampling rate of 9 Hz. For real-time recognition, a sliding window was applied to the data with a window size of 100 and an overlap of 50%. Data augmentation was performed in the same manner as in a previous study [45], and synthetic data were generated to ensure at least 6500 windowed data for each class.

2.3. Training and Evaluation Method

For the evaluation of the proposed algorithm, the data are divided into training, validation, and test datasets in a 4:2:3 ratio. The predictor pre-trains with training data. At this time, the batch-size training data is received from the DVE. The batch size of the predictor is considered as 64 for BBS data and 1024 for public data based on excellent results obtained and corresponding to the batch sizes [35,37]. The predictor uses the Adam optimizer, a learning rate of 0.01, and 200 iterations. DVE uses the reward received from the predictor for training. The DVE uses the Adam optimizer, a learning rate of 0.01, and 30 iterations for training. Algorithm 1 describes the training process of the data valuation algorithm using a pseudocode.

Corrupted sample discovery (CSD) and remove high/low-value samples (RHLVS) are used to evaluate the data valuation algorithm. In total, 20% of the labels of the training data are contaminated for evaluation. CSD represents the rate at which corrupted samples are discovered while accumulating a constant rate of the amount of data from the lowest value data. The performance of the data valuation algorithm can be considered to be excellent when several corrupted samples are observed in the low-value data. In the previous studies on data valuation algorithms [26,30], CSD was used as a performance evaluation criterion. A value is assigned to each sample of the training dataset using the trained data valuation algorithm, and the dataset is in descending order in terms of the value of $D^{sortH} = \left\{ (x_i^h, y_i^h) \right\}_{i=1}^N$. The contaminated training dataset is denoted by $D^{conta} = \{(x, y)\}_{i=1}^0$. The number of data corresponding to 5% of the training data is $v = \frac{N}{100/5}$, and the index of the result is $r \in \{0, 1, \dots, 10\}$. Accumulated data with low values are denoted by $D_r^{sortH} = \{(x_i^H, y_i^H)\}_{i=1}^{v \times r}$. The function that outputs the number of elements

in a dataset is denoted by $Numb$. The formula for obtaining CSD and the resulting value can be expressed as follows:

$$\begin{aligned} CSD_r &= \frac{Numb(D_r^{sortH} \cap D^{conta})}{o} \\ CSD &= \{CSD_0, \dots, CSD_{10}\} \end{aligned} \quad (3)$$

Algorithm 1 Pseudo-code of data valuation training

Inputs: Learning rates $\alpha, \beta > 0$, mini-batch size $B_p, B_s > 0$, inner iteration count $N_I > 0$, moving average window $T > 0$, training dataset D , validation dataset $D^v = \{(x_k^v, y_k^v)\}_{k=1}^L$

Initialize parameters θ, φ , moving average $\delta = 0$

While waiting for convergence **perform**

Sample $D_B = (x_j, y_j)_{j=1}^{B_s} \sim D$

for $j = 1, \dots, B_s$ **perform**

Get selection probabilities : $w_j = h_\varphi(x_j, y_j)$

Sample a selection vector : $s_j \sim Ber(w_j)$

for $t = 1, \dots, N_I$ **perform**

Sample $(\tilde{x}_m, \tilde{y}_m, \tilde{s}_m)_{m=1}^{B_p} \sim (x_j, y_m, s_m)_{j=1}^{B_s}$

Update the predictor model :

$$\theta \leftarrow \theta - \frac{\alpha}{B_p} \sum_{m=1}^{B_p} \tilde{s}_m \cdot \nabla_{\theta} L_f(f_{\theta}(\tilde{x}_m), \tilde{y}_m)$$

Update the DVE model:

$$\varphi \leftarrow \varphi - \left[\frac{\beta}{L} \sum_{k=1}^L [L_h(f_{\theta}(x_k^v), y_k^v)] - \delta \right] \cdot \nabla_{\varphi} \log \pi_{\varphi}(D_B, (s_1, \dots, s_b))$$

Update the baseline:

$$\delta \leftarrow \frac{T-1}{T} \delta + \frac{1}{LT} \sum_{k=1}^L [L_h(f_{\varphi}(x_k^v), y_k^v)]$$

RHLVS repeatedly removes a certain amount of data—either of highest or lowest value—and evaluates the accuracies of HAR after removing each training data point. If the data valuation model exhibits excellent performance, the classification accuracy can be decreased by removing high-value data, and the classification accuracy can be slightly improved by removing low-value data. The value (=selection probability) of each training sample calculated using the data valuation algorithm is given by $h_{\varphi}(D) = \{p_1, \dots, p_N\}$. When $Sort_h$ is a descending sort function and $Sort_L$ is an ascending sort function, $Sort_h(h_{\varphi}(D)) = \{p_{11}^h, \dots, p_N^h\}$, $Sort_L(h_{\varphi}(D)) = \{p_1^l, \dots, p_N^l\}$. The sorted training data are denoted by $D^{sortL} = \{(x_i^l, y_i^l)\}_{i=1}^N$ and $D^{sortH} = \{(x_i^h, y_i^h)\}_{i=1}^N$. The classification model is denoted by C_{γ} , where γ denotes the training parameter. The accuracy function, *Accuracy*, calculates the accuracy using prediction labels and true test labels. The formulas and results used to obtain “remove high value data” and “remove low value data” can be expressed as follows:

$$\begin{aligned} C_{\gamma_r}^h &= \operatorname{argmin}_{f \in F} \frac{1}{N-v^*r} \sum_{i=v^*r}^N p_i^h \cdot \nabla_{L_C} (\hat{C}_{\gamma}(\mathbf{x}_i^h), y_i^h) \\ C_{\gamma_r}^l &= \operatorname{argmin}_{f \in F} \frac{1}{N-v^*r} \sum_{i=v^*r}^N p_i^l \cdot \nabla_{L_C} (\hat{C}_{\gamma}(\mathbf{x}_i^l), y_i^l) \\ Acc_r^h &= \operatorname{Accuracy}(C_{\gamma_r}^h(D^t), y^t) \\ Acc_r^l &= \operatorname{Accuracy}(C_{\gamma_r}^l(D^t), y^t) \\ \text{remove high value data} &= [Acc_0^h, \dots, Acc_{10}^h] \\ \text{remove low value data} &= [Acc_0^l, \dots, Acc_{10}^l] \end{aligned} \quad (4)$$

For RHLVS, a deep learning-based classification model is used. A model structurally identical to the model for BBS HAR data proposed in [35] but with a slightly lower capacity is used for the BBS HAR data. The complex model for public HAR data proposed in [36] is

used for public data, and the baseline model for public HAR data is used for comparison. Figure 4 depicts the structure of the HAR model for BBS and public data. The 1D-CNN convolutional layer used in the three models has a filter size = 64, kernel size = 3, activation function = “ReLU”, and maxpooling size = 2. The unit size of the GRU layer is taken to be eight, and an 8-size hidden state is the output for each unit. The Model for BBS HAR data uses a 50% dropout layer, and the complex model for public HAR data uses a 70% dropout layer. The number of perceptrons in the dense layer is 100 in all three models.

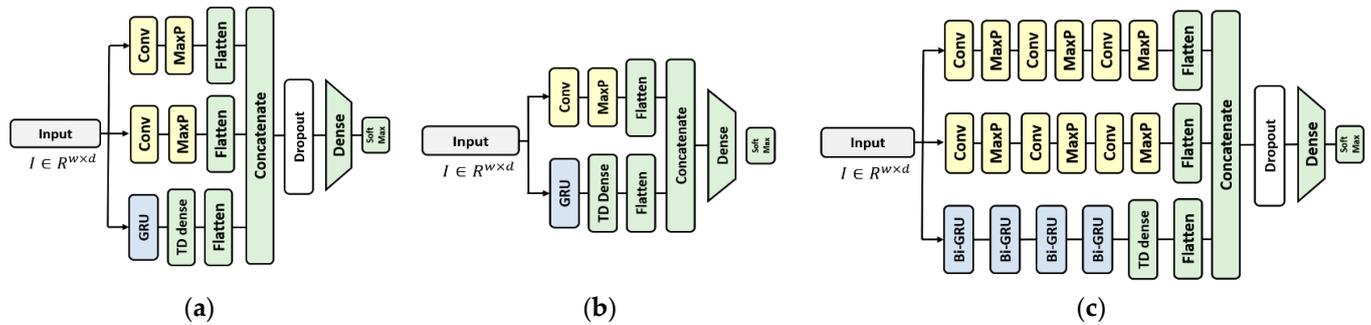


Figure 4. Deep learning-based HAR model for performance evaluation of the data valuation algorithm: (a) the model for BBS HAR data, (b) baseline model for public HAR data, and (c) complex model for public HAR data.

3. Results and Discussion

3.1. Evaluation of the Proposed Algorithm on BBS HAR Data

The performance of the data valuation algorithm is evaluated through CSD and RHLVS after training the algorithm using the BBS HAR data. Figures 5–11 depict the CSD and RHLVS of the proposed algorithm on the BBS HAR data. In CSD, the maximum corrupted discovery rate is plotted on a graph. In RHLVS, the maximum and minimum accuracies are indicated on the graph.

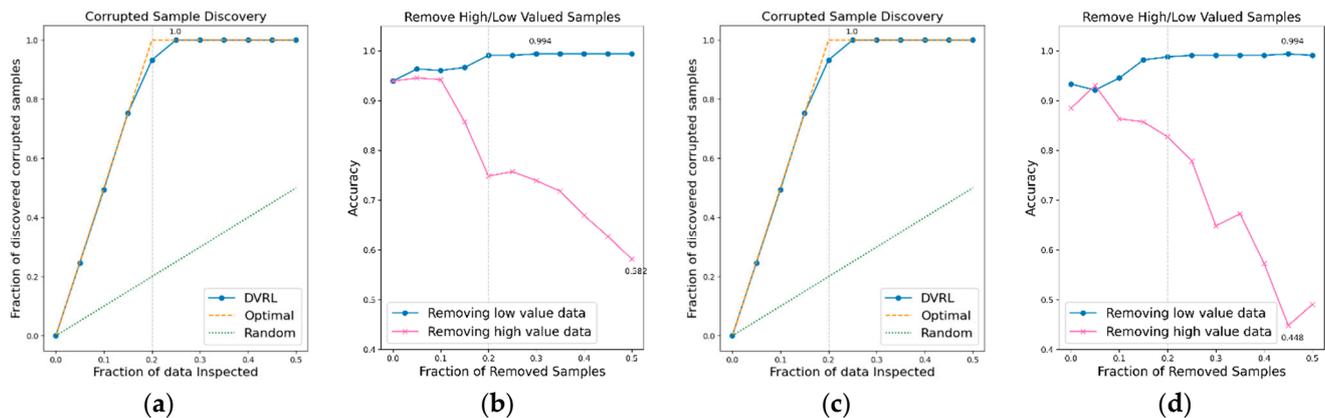


Figure 5. CSD and RHLVS on BBS HAR data: (a) CSD graph of BBS task 1, (b) RHLVS graph of BBS task 1, (c) CSD graph of BBS task 2, and (d) RHLVS graph of BBS task 2.

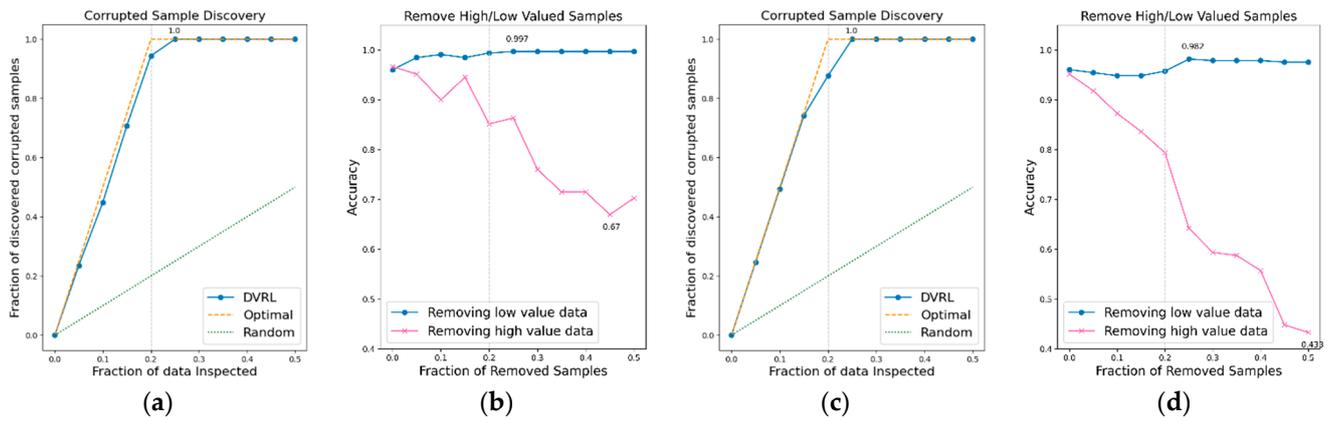


Figure 6. CSD and RHLVS on BBS HAR data: (a) CSD graph of BBS task 3, (b) RHLVS graph of BBS task 3, (c) CSD graph of BBS task 4, and (d) RHLVS graph of BBS task 4.

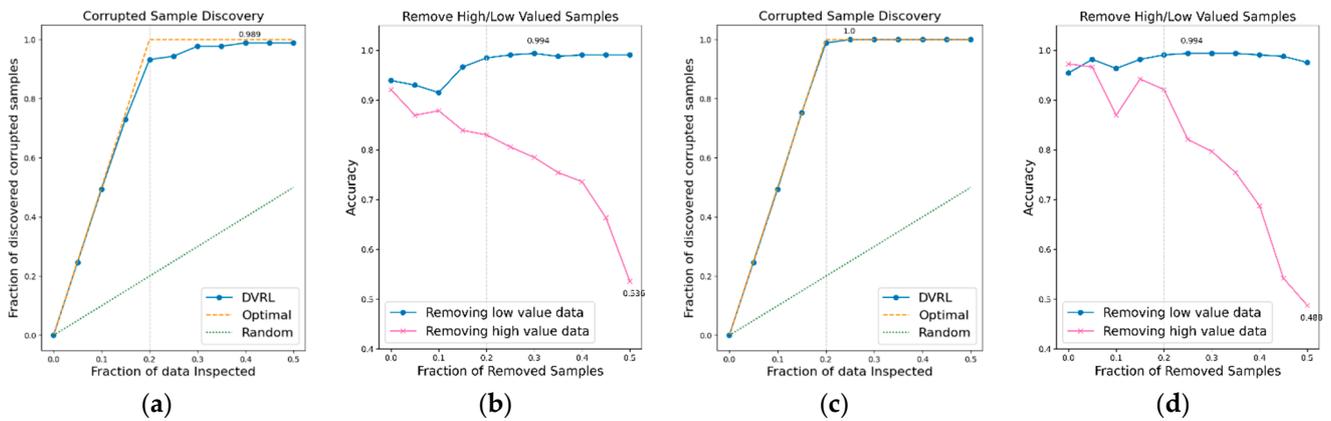


Figure 7. CSD and RHLVS on BBS HAR data: (a) CSD graph of BBS task 5, (b) RHLVS graph of BBS task 5, (c) CSD graph of BBS task 6, and (d) RHLVS graph of BBS task 6.

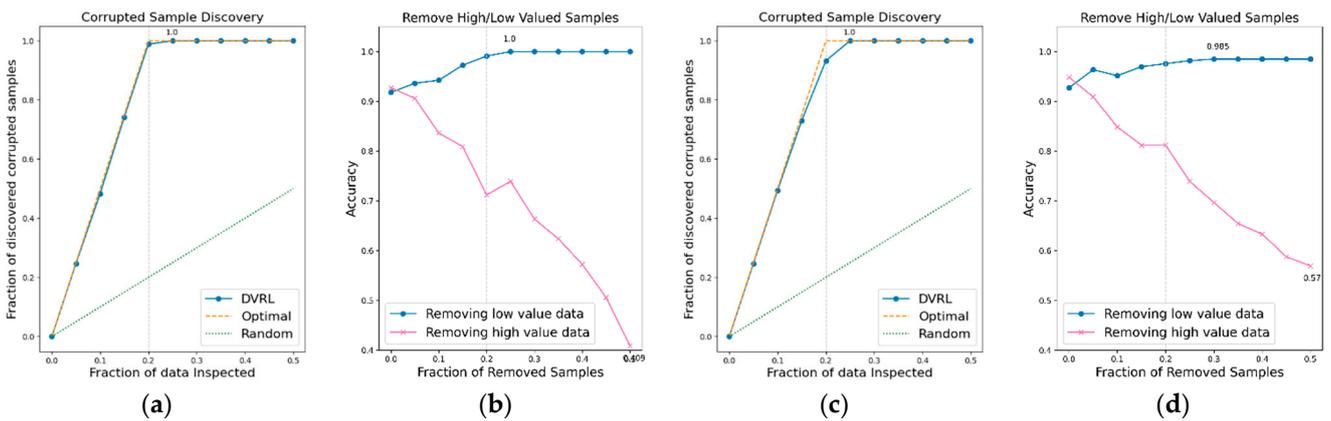


Figure 8. CSD and RHLVS on BBS HAR data: (a) CSD graph of BBS task 7, (b) RHLVS graph of BBS task 7, (c) CSD graph of BBS task 8, and (d) RHLVS graph of BBS task 8.

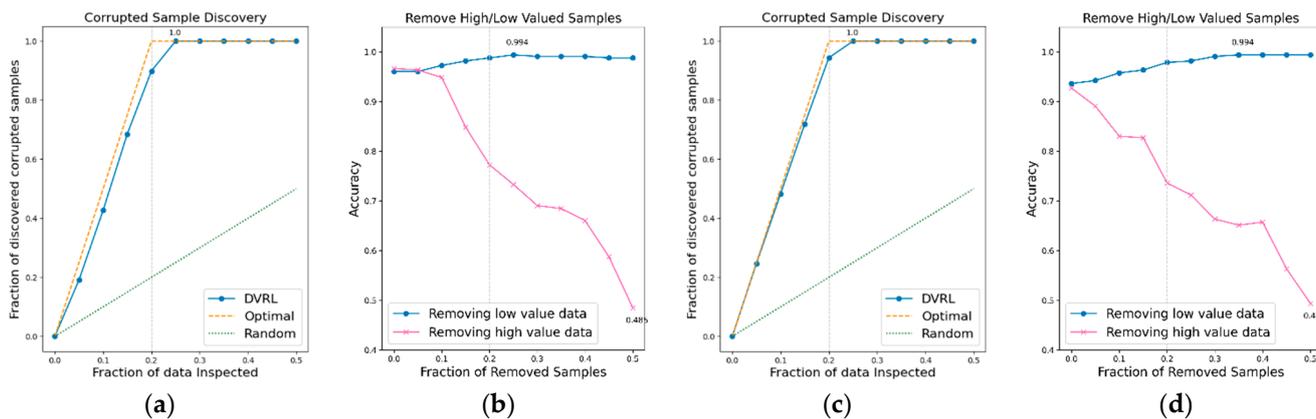


Figure 9. CSD and RHLVS on BBS HAR data: (a) CSD graph of BBS task 9, (b) RHLVS graph of BBS task 9, (c) CSD graph of BBS task 10, and (d) RHLVS graph of BBS task 10.

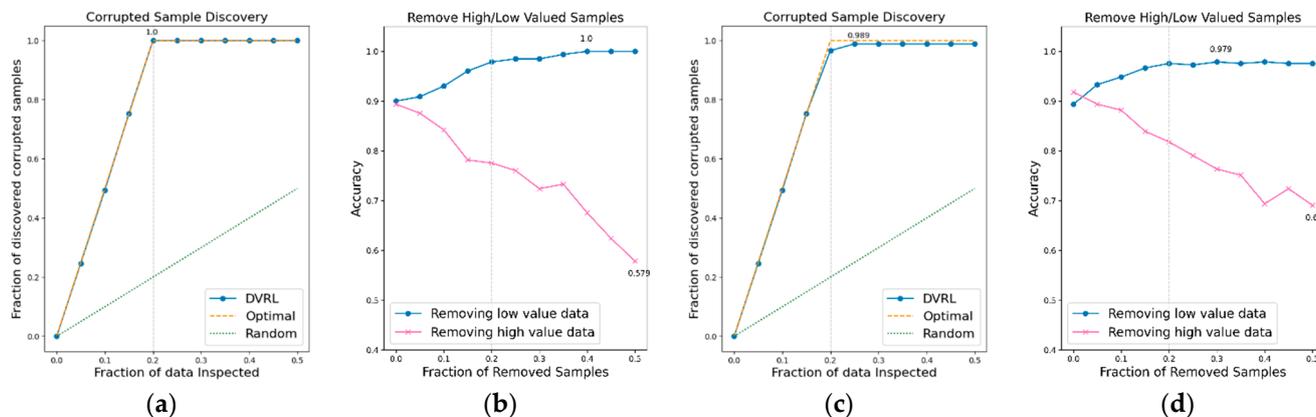


Figure 10. CSD and RHLVS on BBS HAR data: (a) CSD graph of BBS task 11, (b) RHLVS graph of BBS task 11, (c) CSD graph of BBS task 12, and (d) RHLVS graph of BBS task 12.

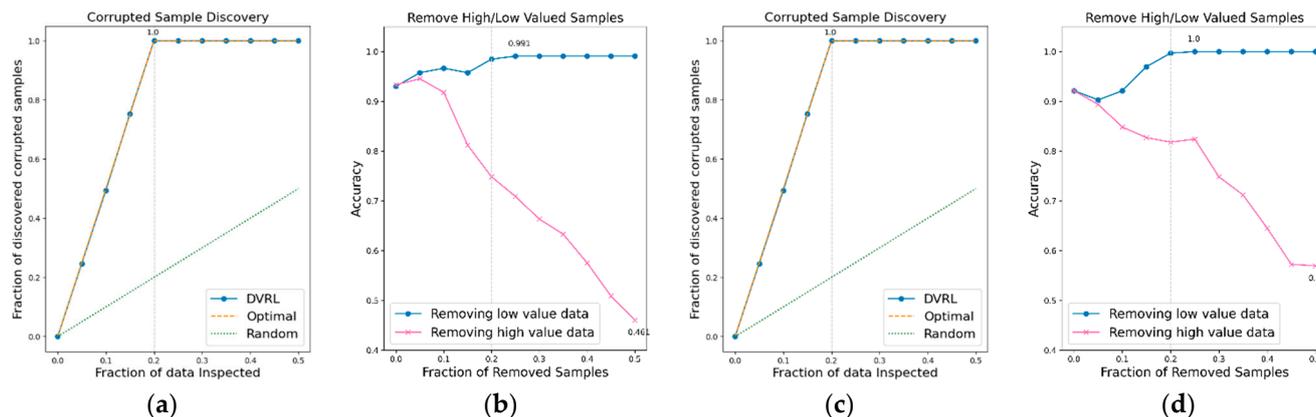


Figure 11. CSD and RHLVS on BBS HAR data: (a) CSD graph of BBS task 13, (b) RHLVS graph of BBS task 13, (c) CSD graph of BBS task 14, and (d) RHLVS graph of BBS task 14.

In the CSD graphs depicted in Figures 5a,c, 6a,c, 7a,c, 8a,c, 9a,c, 10a,c and 11a,c, the optimal graph corresponding to ideal model performance is drawn. Greater amounts of contaminated data in the low-value data correspond to higher proximity between the algorithm’s performance graph and the optimal graph. All data graphs depicted in BBS tasks 1–14 are confirmed to be close to the optimal graph, corroborating the excellent performance of the algorithm. The RHLVS graphs depicted in Figures 5b,d, 6b,d, 7b,d, 8b,d, 9b,d, 10b,d and 11b,d

indicate that, in the remove high value data graph, the accuracy decreases significantly as the percentage of data to be removed increases. Meanwhile, in the remove low value data graph, the reduction in accuracy is small even when the percentage of data to be removed is high. As this tendency is clear, the performance of the data valuation algorithm is considered to be good. The performance is improved by removing low-value data from all BBS data. Thus, the data valuation algorithm improves classification performance by improving the training data quality.

Table 1 summarizes the major results obtained for RHLVS and CSD on BBS data, including maximum accuracy, improved accuracy, and removed data. Improved accuracy indicates the maximum improved performance achieved by removing low-value data. Removed data represents the rate of removal of low-value data at maximum accuracy. For CSD, the major metrics are maximum discovery and removed data. Maximum discovery represents the ratio of the corrupted data when the discovery of corrupted data is the maximum. Removed data represents the ratio of removed low-value data when the discovery of corrupted data is the maximum.

Table 1. Major metrics of data valuation algorithm on BBS data.

Task	Major Figures of RHLVS (%)			Major Figures of CSD (%)	
	Maximum Accuracy	Improved Accuracy	Removed Data	Maximum Discovery	Removed Data
1	99.4	5.5	30	100	25
2	99.4	6.1	45	100	25
3	99.7	3.6	25	100	25
4	98.2	2.1	25	100	25
5	99.4	5.5	30	98.9	40
6	99.4	3.9	25	100	25
7	100	8.2	25	100	25
8	98.5	5.8	30	100	25
9	99.4	3.3	25	100	25
10	99.4	5.8	35	100	25
11	100	10	40	100	20
12	97.9	8.5	30	98.9	25
13	99.1	6.1	25	100	20
14	100	7.9	25	100	20
Average	99.3	5.9	29.5	99.8	25

The CSD values on BBS tasks 1–14 indicate that the average maximum discovery is 99.8%, and the average removed data is 25%. As the corrupted data comprises 20% of the training data, the CSD performance is almost ideal. The RHLVS values indicate that the average maximum accuracy is 99.3%, which corresponds to excellent performance, and the average improved accuracy is 5.9%. Furthermore, a performance improvement is confirmed in all the tasks. Data with values below 25% are observed to be primarily composed of contaminated data. Therefore, the performance improvement observed in the initial part of the remove low value data graph seems to be primarily caused by the removal of contaminated data. As the average removed data for RHLVS is 29.5% and the that for CSD is 25%, data with a low value among the clean data are also removed in the former case.

3.2. Evaluation of the Proposed Algorithm on BBS HAR Data

The primary purpose of the data valuation algorithm is to improve classification performance by enhancing the quality of training data. In this study, performance improvement is confirmed by removing low-value data from the training data in RHLVS using the proposed data valuation algorithm. A comparison of the results of this study with those of a previous study on BBS HAR [35] reveals the extent of improvement. This study uses the same data and a structurally identical model with a slightly smaller capacity as [35].

Table 2 depicts the performance reported in [35], the maximum accuracy calculated from the RHLVS in this study, and the volume of training data used for training. In this study, 44% of the total data are used as training data, which is determined by accounting for the ratio of the removed data at the maximum accuracy for RHLVS.

Table 2. Comparison of BBS scoring performances of [35] and after applying the DVRL algorithm.

Task	Accuracy (%)		Amount of Training Data (%)	
	Previous Study [35]	This Study	Previous Study [35]	This Study
1	98.5	99.4	90	30.8
2	98.5	99.4	90	24.2
3	99.6	99.7	90	33
4	99	98.2	90	33
5	96.7	99.4	90	30.8
6	97.9	99.4	90	33
7	99	100	90	33
8	98.9	98.5	90	30.8
9	97.8	99.4	90	33
10	98.2	99.4	90	28.6
11	97.8	100	90	26.4
12	98.2	97.9	90	30.8
13	98.1	99.1	90	33
14	99.1	100	90	33
Average	98.4	99.3	90	31

By improving the corrupted BBS HAR data by applying the data valuation algorithm, the proposed algorithm outperformed the method proposed in [35], which used clean data. The application of the proposed data valuation algorithm improves performance perceptibly as corrupted training data as well as low-value data are removed from the clean data. Moreover, excellent performance is confirmed when a small quantity of high-quality data is used—the algorithm proposed in this study uses approximately 59% less training data on average than that of [35].

3.3. Evaluation of the Proposed Algorithm on Public HAR Data

An additional experiment is conducted to verify if the proposed data valuation algorithm, which exhibits good performance on BBS HAR data, continues to perform well on public IMU-based HAR data. The data valuation algorithm is applied to public IMU-based HAR datasets, UCI-HAR, WISDM, and PAMAP2, and its performance is evaluated in terms of CSD and RHLVS. The baseline model for public HAR data and complex model for public HAR data are used to evaluate the performance of the data valuation algorithm on public IMU-based HAR data. Figures 12–14 depict the CSD and RHLVS results on public HAR data. In the case of CSD, the maximum corrupted discovery rate is plotted on the graph. In the case of RHLVS, the maximum and minimum accuracies are indicated on the graph and the baseline, and complex model for public HAR data are used for classification.

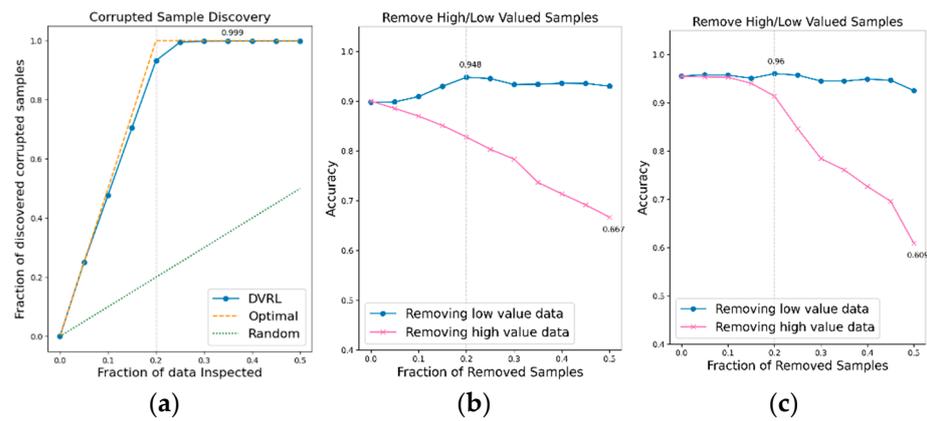


Figure 12. CSD and RHLVS on IMU-based HAR public data: (a) CSD graph of UCI-HAR data, (b) RHLVS graph of UCI-HAR data using baseline model for public HAR data, and (c) RHLVS graph of UCI-HAR data using the complex model for public HAR data.

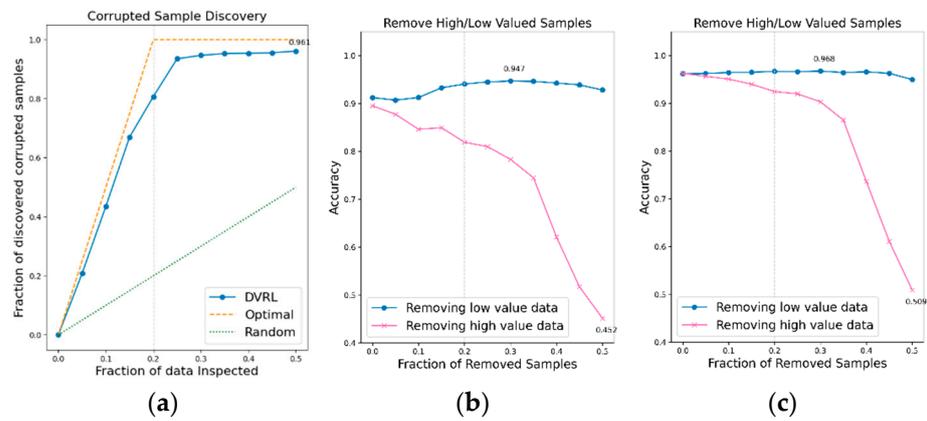


Figure 13. CSD and RHLVS on IMU-based HAR public data: (a) CSD graph of WISDM data, (b) RHLVS graph of WISDM data using baseline model for public HAR data, and (c) RHLVS graph of WISDM data using the complex model for public HAR data.

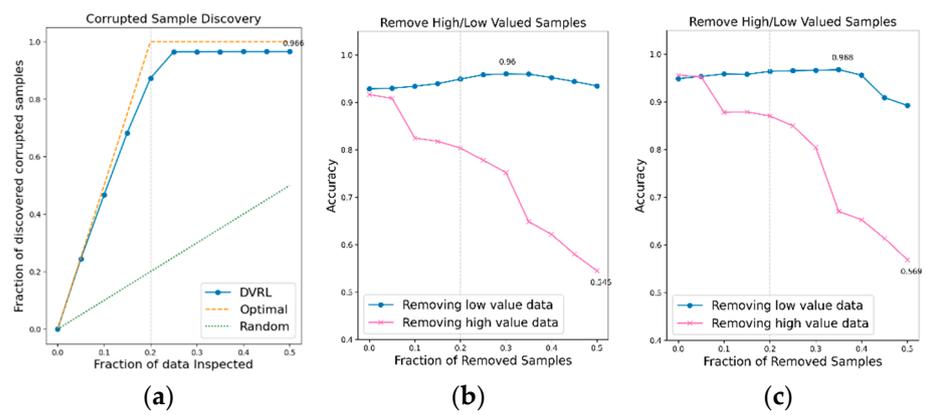


Figure 14. CSD and RHLVS on IMU-based HAR public data: (a) CSD graph of PAMAP2 data, (b) RHLVS graph of PAMAP2 data using baseline model for public HAR data, and (c) RHLVS graph of PAMAP2 data using the complex model for public HAR data.

Figures 12a,c, 13a,c and 14a,c indicate that the CSD graphs are close to the optimal graphs on all public HAR datasets, and the performance of the algorithm is good in all cases, despite being inferior to that of BBS data. Figures 12b,d, 13b,d and 14b,d indicate that the accuracy of the removed high-value data graph decreases significantly as the

percentage of data removed increases. On the other hand, in the removed low-value data graph, the reduction in accuracy is small even when the percentage of data removed is high. As this tendency is clear, the data valuation algorithm can be considered to perform well. In the removed low-value data graph, the performance improvement is attributed to the removal of low-value data. Thus, the performance of the classification model is improved by improving the quality of training data using the proposed data valuation algorithm.

Table 3 presents the major metrics of RHLVS and CSD. For RHLVS, these are the maximum accuracy, improved accuracy, and removed data. Improved accuracy represents the maximum improved performance while removing low-value data. The removed data represents the rate of removal of low-value data at maximum accuracy. The major metrics of CSD are maximum discovery and removed data. The maximum discovery represents the ratio of corrupted data when the discovery of corrupted data is at its maximum. The removed data represents the ratio of the removed low-value data when the discovery of corrupted data is at its maximum.

Table 3. Major metrics used to evaluate the proposed data valuation algorithm on public IMU-based HAR data.

Data	Major Metrics of RHLVS Baseline Model for Public HAR Data (%)			Major Metrics of RHLVS Complex Model for Public HAR Data (%)			Major Metrics of CSD (%)	
	Maximum Accuracy	Improved Accuracy	Removed Data	Maximum Accuracy	Improved Accuracy	Removed Data	Maximum Discovery	Removed Data
UCI-HAR	94.8	5	20	96.0	0.4	20	99.9	35
WISDM	94.7	3.5	30	96.8	0.6	30	96.1	50
PAMAP2	96.0	3.1	30	96.8	1.9	35	96.6	50

The CSD on the IMU-based HAR data reveals that more than 96% of the contaminated data is identified on all IMU-based HAR datasets. Maximum discovery and the removed data of CSD are observed to be 35%, 50%, and 50% of the UCI-HAR, WISDM, and PAMAP2 data, respectively. However, when 25% of the low-value data are accumulated in all three datasets, the discovery rate of corrupted data becomes close to the maximum discovery rate. Considering that corrupted data accounts for 20% of the training data, the CSD performance can be considered to be excellent. In RHLVS, the maximum accuracies in the baseline model for public HAR data are 94.8%, 94.7%, and 96.0% for the UCI-HAR, WISDM, and PAMAP2 datasets, respectively, and 96.0%, 96.8%, and 96.8% for the complex model for public HAR data. The accuracy is improved by 5%, 3.5%, and 3.1% on the UCI-HAR, WISDM, and PAMAP2 datasets, respectively, when using the Baseline model for public HAR data, and by 0.4%, 0.6%, and 1.9%, respectively, when using the Complex model for public HAR data. Performance improvement is confirmed in all experiments by improving the data quality using the data valuation algorithm. The improvement in accuracy over the baseline model for public HAR data is greater than that over the complex model for public HAR data owing to the better regularization performance of the latter. In RHLVS, the removed data value is 20%, 30%, and 50% for UCI-HAR, WISDM, and PAMAP2 data, respectively. When 25% of the low-value data is removed, the CSD performance is observed to be almost saturated on all three datasets. As the number of training data gradually decreases as the data are removed, it seems that the maximum accuracy is attained before reaching the maximum discovery. In conclusion, the proposed data-valuation algorithm is observed to exhibit excellent classification performance on IMU-based HAR public data.

4. Conclusions

In this paper, a meta-reinforcement learning-based data-valuation algorithm was proposed to improve the IMU-based HAR training data. A deep learning structure suitable for IMU-based HAR was introduced in the DVE, and a predictor was added to construct

the data valuation algorithm. In previous studies [26,29,30], vision-related public datasets were used primarily, because the purpose of the data valuation algorithm was limited to improve image classification. In this study, HAR data, which is multivariate timeseries data derived from human movements, is targeted. In general, HAR data has a lower resolution than vision data, and the amount of data in a public dataset is relatively very small. The purpose of our study is to find a good data valuation algorithm that functions efficiently on HAR data. The proposed data valuation algorithm improves the DVE structure used in previous studies [30], which were not capable of feature extraction after adding this capability. Therefore, the proposed algorithm can accept raw-level data as an input, making pre-training redundant. In previous studies, a pre-trained model was used to train the network for prediction purposes. However, the feature extraction network of the proposed algorithm was trained explicitly for data valuation. As the feature extraction structure, a transformer encoder block was inserted after the 1D-CNN backbone in front of the DVE, enabling the extraction of both local and global features. The proposed data-valuation algorithm is observed to be capable of performing feature extraction and data-value estimation simultaneously. The predictor uses a multi-head 1D-CNN-based stacking ensemble structure with good efficiency and performance on IMU-based HAR data. Two metrics, CSD and RHLVS, were used to evaluate the algorithm. In terms of CSD, the ability of the algorithm to discover corrupted data is observed to be excellent on all four IMU-based HAR datasets. In particular, the ability to identify corrupted BBS HAR data is nearly ideal. In terms of RHLVS, the performance is observed to be improved by removing low-value data. The proposed data valuation algorithm exhibits excellent performance in assigning data values to all IMU-based HAR data, confirming that it can contribute to the improvement of the quality of IMU-based HAR data and HAR model performance.

The proposed data valuation algorithm suffers from the limitation of requiring manual updating of clean validation data when new data are added, which can be time-consuming and effort-intensive. We intend to compensate for this limitation in a follow-up study by utilizing a semi-supervised learning algorithm to remove outliers and identify good-quality validation data. Another limitation of the proposed data valuation algorithm is that its entire structure cannot use gradient descent or backpropagation, which is commonly used for training deep neural network algorithms, as its sampler structure is non-differentiable. This is why the algorithm uses meta-reinforcement learning instead. In a follow-up study, we intend to use an alternative differentiable structure, enabling the algorithm to be trained using gradient descent or backpropagation. This may improve training efficiency and performance.

Recently, large technology companies have started to gather smartphone and smart watch-based healthcare data automatically to provide more comprehensive healthcare service. These data require management and quality control to ensure good service. We expect the proposed algorithm to be efficient and effective in this regard.

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

Funding: This study was supported by Inha University and “Cooperative Research Program for Agriculture Science and Technology Development (Project No. PJ01531203)” Rural Development Administration, Republic of Korea.

Institutional Review Board Statement: The study was conducted according to the guidelines of the Declaration of Helsinki and approved by the Institutional Review Board of INHA UNIVERSITY HOSPITAL, Incheon, Korea (IRB license number 2016-08-012 and date of approval 10 December 2016).

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The experiments have been carried out using sensor-based HAR datasets such as UCI [42], WISDM [43] and PAMAP2 [44] which are open for use in the research work.

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

References

1. Bayouhdh, K.; Knani, R.; Hamdaoui, F.; Mtibaa, A. A survey on deep multimodal learning for computer vision: Advances, trends, applications, and datasets. *Vis. Comput.* **2022**, *38*, 2939–2970. [[CrossRef](#)] [[PubMed](#)]
2. Latif, S.; Cuayahuitl, H.; Pervez, F.; Shamshad, F.; Ali, H.S.; Cambria, E. A survey on deep reinforcement learning for audio-based applications. *Artif. Intell. Rev.* **2022**, 1–48. [[CrossRef](#)]
3. Lauriola, I.; Lavelli, A.; Aioli, F. An introduction to deep learning in natural language processing: Models, techniques, and tools. *Neurocomputing* **2022**, *470*, 443–456. [[CrossRef](#)]
4. Li, Y.; Huang, C.; Ding, L.; Li, Z.; Pan, Y.; Gao, X. Deep learning in bioinformatics: Introduction, application, and perspective in the big data era. *Methods* **2019**, *166*, 4–21. [[CrossRef](#)] [[PubMed](#)]
5. Yu, Y.; Li, M.; Liu, L.; Li, Y.; Wang, J. Clinical big data and deep learning: Applications, challenges, and future outlooks. *Big Data Min. Anal.* **2019**, *2*, 288–305. [[CrossRef](#)]
6. Tariq, M.I.; Tayyaba, S.; Ashraf, M.W.; Balas, V.E. Deep learning techniques for optimizing medical big data. In *Deep Learning Techniques for Biomedical and Health Informatics*; Elsevier: Amsterdam, The Netherlands, 2020; pp. 187–211.
7. Dang, H.V.; Tran-Ngoc, H.; Nguyen, T.V.; Bui-Tien, T.; De Roeck, G.; Nguyen, H.X. Data-driven structural health monitoring using feature fusion and hybrid deep learning. *IEEE Trans. Autom. Sci. Eng.* **2020**, *18*, 2087–2103. [[CrossRef](#)]
8. Rueckert, D.; Schnabel, J.A. Model-based and data-driven strategies in medical image computing. *Proc. IEEE* **2019**, *108*, 110–124. [[CrossRef](#)]
9. Ma, X.; Gao, Z. Data-driven deep learning to design pilot and channel estimator for massive MIMO. *IEEE Trans. Veh. Technol.* **2020**, *69*, 5677–5682. [[CrossRef](#)]
10. Lee, H.; Kim, Y.; Kim, C.O. A deep learning model for robust wafer fault monitoring with sensor measurement noise. *IEEE Trans. Semicond. Manuf.* **2016**, *30*, 23–31. [[CrossRef](#)]
11. Siamwala, A.; Lochhead, Z.; Abdulla, W. Environmental noise monitoring using distributed IoT sensor nodes. In Proceedings of the International Conference on Electronics, Information, and Communication (ICEIC), Auckland, New Zealand, 22–25 January 2019; IEEE: New York, NY, USA, 2019; pp. 1–10.
12. Xia, S.; Chen, B.; Wang, G.; Zheng, Y.; Gao, X.; Giem, E.; Chen, Z. mCRF and mRD: Two classification methods based on a novel multiclass label noise filtering learning framework. *IEEE Trans. Neural Netw. Learn. Syst.* **2021**, *33*, 2916–2930. [[CrossRef](#)]
13. Sun, L.; Wang, T.; Ding, W.; Xu, J.; Lin, Y. Feature selection using Fisher score and multilabel neighborhood rough sets for multilabel classification. *Inf. Sci.* **2021**, *578*, 887–912. [[CrossRef](#)]
14. Zhang, C.; Bengio, S.; Hardt, M.; Recht, B.; Vinyals, O. Understanding deep learning requires rethinking generalization. In Proceedings of the 5th International Conference on Learning Representations, Toulon, France, 24–26 April 2017.
15. Xiao, T.; Xia, T.; Yang, Y.; Huang, C.; Wang, X. Learning from massive noisy labeled data for image classification. In Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition, Boston, MA, USA, 7–12 June 2015; IEEE: New York, NY, USA, 2015; pp. 2691–2699.
16. Li, S.Y.; Huang, S.J.; Chen, S. Crowdsourcing aggregation with deep Bayesian learning. *Sci. China Inf. Sci.* **2021**, *64*, 11. [[CrossRef](#)]
17. Zhu, X.; Wu, X. Class noise vs. attribute noise: A quantitative study. *Artif. Intell. Rev.* **2004**, *22*, 177–210. [[CrossRef](#)]
18. Peng, X.; Wang, P.; Xia, S.; Wang, C.; Chen, W. VPGb: A granular-ball based model for attribute reduction and classification with label noise. *Inf. Sci.* **2022**, *611*, 504–521.
19. Lu, Y.; Bo, Y.; He, W. An Ensemble model for combating label noise. In Proceedings of the Fifteenth ACM International Conference on Web Search and Data Mining, Virtual Event, AZ, USA, 21–25 February 2022; ACM: New York, NY, USA, 2022; pp. 608–617.
20. Kuang, M.; Wang, W.; Chen, Z.; Kang, L.; Yan, Q. Efficient two-stage label noise reduction for retrieval-based tasks. In Proceedings of the Fifteenth ACM International Conference on Web Search and Data Mining, Virtual Event, AZ, USA, 21–25 February 2022; ACM: New York, NY, USA, 2022; pp. 526–534.
21. Liu, Y.; Yao, R.; Jia, S.; Wang, F.; Wang, R.; Ma, R.; Qi, L. A label noise filtering and label missing supplement framework based on game theory. *Digit. Commun. Netw.* **2022**, in press. [[CrossRef](#)]
22. Shapley, L.S. A value for n-person games. *Contrib. Theory Games* **1953**, *2*, 307–317.
23. Ju, L.; Wang, X.; Wang, L.; Mahapatra, D.; Zhao, X.; Zhou, Q.; Liu, T.; Ge, Z. Improving Medical Images Classification with Label Noise Using Dual-Uncertainty Estimation. *IEEE Trans. Med. Imaging* **2022**, *41*, 1533–1546. [[CrossRef](#)]
24. Shapiro, A. Monte Carlo sampling methods. In *Handbooks in Operations Research and Management Science*; Elsevier: Amsterdam, The Netherlands, 2003; Volume 10, pp. 353–425.
25. Cook, R.D. Detection of influential observation in linear regression. *Technometrics* **1977**, *19*, 15–18.
26. Ghorbani, A.; Zou, J. Data shapley: Equitable valuation of data for machine learning. In Proceedings of the International Conference on Machine Learning, Long Beach, CA, USA, 9–15 June 2019; pp. 2242–2251.
27. Ren, M.; Zeng, W.; Yang, B.; Urtasun, R. Learning to reweight examples for robust deep learning. In Proceedings of the International Conference on Machine Learning, Stockholm, Sweden, 10–15 July 2018; PMLR; pp. 4334–4343.

28. Hendrycks, D.; Mazeika, M.; Wilson, D.; Gimpel, K. Using trusted data to train deep networks on labels corrupted by severe noise. In Proceedings of the Advances in Neural Information Processing Systems 2018, Montreal, QC, Canada, 3–8 December 2018; p. 31.
29. Saeed, S.U.; Fu, Y.; Stavrinides, V.; Baum, Z.M.; Yang, Q.; Rusu, M.; Fan, R.E.; Sonn, G.A.; Noble, J.A.; Barratt, D.C.; et al. Image quality assessment for machine learning tasks using meta-reinforcement learning. *Med. Image Anal.* **2022**, *78*, 102427.
30. Yoon, J.; Arik, S.; Pfister, T. Data valuation using reinforcement learning. In Proceedings of the International Conference on Machine Learning, Virtual, 13–18 July 2020; pp. 10842–10851.
31. Mekruksavanich, S.; Jitpattanakul, A. Smartwatch-based human activity recognition using hybrid lstm network. In Proceedings of the 2020 IEEE Sensors, Virtual, 25–28 October 2020; IEEE: New York, NY, USA, 2020; pp. 1–4.
32. Mukherjee, D.; Mondal, R.; Singh, P.K.; Sarkar, R.; Bhattacharjee, D. EnsemConvNet: A deep learning approach for human activity recognition using smartphone sensors for healthcare applications. *Multimed. Tools Appl.* **2020**, *79*, 31663–31690. [[CrossRef](#)]
33. Ordóñez, F.J.; Roggen, D. Deep convolutional and lstm recurrent neural networks for multimodal wearable activity recognition. *Sensors* **2016**, *16*, 115. [[CrossRef](#)] [[PubMed](#)]
34. Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A.N.; Kaiser, L.; Polosukhin, I. Attention is all you need. In Proceedings of the Advances in Neural Information Processing Systems 2017, Long Beach, CA, USA, 4–9 December 2017; NIPS: Long Beach, CA, USA, 2018; p. 30.
35. Kim, Y.W.; Joa, K.L.; Jeong, H.Y.; Lee, S. Wearable IMU-based human activity recognition algorithm for clinical balance assessment using 1D-CNN and GRU ensemble model. *Sensors* **2021**, *21*, 7628. [[CrossRef](#)] [[PubMed](#)]
36. Gholamrezaii, M.; AlModarresi, S.M.T. A time-efficient convolutional neural network model in human activity recognition. *Multimed. Tools Appl.* **2021**, *80*, 19361–19376. [[CrossRef](#)]
37. Kim, Y.W.; Cho, W.H.; Kim, K.S.; Lee, S. Oversampling technique-based data augmentation and 1D-CNN and bidirectional GRU ensemble model for human activity recognition. *J. Mech. Med. Biol.* **2022**, *22*, 2240048. [[CrossRef](#)]
38. Banjarey, K.; Sahu, S.P.; Dewangan, D.K. Human activity recognition using 1D convolutional neural network. In *Sentimental Analysis and Deep Learning*; Springer: Singapore, Singapore, 2022; pp. 691–702.
39. Berg, K.; Wood-Dauphine, S.; Williams, J.I.; Gayton, D. Measuring balance in the elderly: Preliminary development of an instrument. *Physiother. Can.* **1989**, *41*, 304–311. [[CrossRef](#)]
40. Berg, K.; Wood-Dauphinee, S.; Williams, J.I. The Balance Scale: Reliability assessment with elderly residents and patients with an acute stroke. *Scand. J. Rehabil. Med.* **1995**, *27*, 27–36.
41. Elliott, J. The clinical uses of the Berg Balance Scale. *Physiotherapy* **1997**, *7*, 363. [[CrossRef](#)]
42. Anguita, D.; Ghio, A.; Oneto, L.; Parra Perez, X.; Reyes Ortiz, J.L. A public domain dataset for human activity recognition using smartphones. In Proceedings of the 21th international European symposium on artificial neural networks, computational intelligence and machine learning, Bruges, Belgium, 24–26 April 2013; pp. 437–442.
43. Kwapisz, J.R.; Weiss, G.M.; Moore, S.A. Activity recognition using cell phone accelerometers. *ACM SigKDD Explor. Newsl.* **2011**, *12*, 74–82. [[CrossRef](#)]
44. Reiss, A.; Stricker, D. Introducing a new benchmarked dataset for activity monitoring. In Proceedings of the 2012 16th International Symposium on Wearable Computers, Newcastle, UK, 18–22 June 2012; IEEE: New York, NY, USA, 2012; pp. 108–109.
45. Kim, Y.W.; Cho, W.H.; Kim, K.S.; Lee, S. Inertial-Measurement-Unit-Based Novel Human Activity Recognition Algorithm Using Conformer. *Sensors* **2022**, *22*, 3932. [[CrossRef](#)]

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