electronics

MDPI

Article

A Convolutional Neural Network-Based Feature Extraction and Weighted Twin Support Vector Machine Algorithm for Context-Aware Human Activity Recognition

Kwok Tai Chui [1,*], Brij B. Gupta [2,3,4,5,*], Miguel Torres-Ruiz [6], Varsha Arya [7,8], Wadee Alhalabi [8] and Ikhlas Fuad Zamzami [9]

1 Department of Electronic Engineering and Computer Science, School of Science and Technology, Hong Kong Metropolitan University, Hong Kong, China
2 International Center for AI and Cyber Security Research and Innovations, Department of Computer Science and Information Engineering, Asia University, Taichung 41354, Taiwan
3 School of Information Technology, Skyline University College, Sharjah P.O. Box 1797, United Arab Emirates
4 Center for Interdisciplinary Research, University of Petroleum and Energy Studies (UPES), Dehradun 248007, India
5 University Center for Research & Development (UCRD), Chandigarh University, Chandigarh 140413, India
6 Centro de Investigación en Computación (CIC), Instituto Politécnico Nacional (IPN), Mexico City 07738, Mexico
7 Department of Business Administration, Asia University, Taichung 41354, Taiwan
8 Department of Computer Science, Immersive Virtual Reality Research Group, King Abdulaziz University, Jeddah 21589, Saudi Arabia
9 Management Information System Department, Business College, King Abdulaziz University, Rabigh 21589, Saudi Arabia
* Correspondence: jktchui@hkmu.edu.hk (K.T.C.); bbgupta.nitkkr@gmail.com (B.B.G.)

Citation: Chui, K.T.; Gupta, B.B.; Torres-Ruiz, M.; Arya, V.; Alhalabi, W.; Zamzami, I.F. A Convolutional Neural Network-Based Feature Extraction and Weighted Twin Support Vector Machine Algorithm for Context-Aware Human Activity Recognition. *Electronics* **2023**, *12*, 1915. https://doi.org/10.3390/electronics12081915

Academic Editor: Young-Joo Suh

Received: 6 February 2023
Revised: 29 March 2023
Accepted: 11 April 2023
Published: 18 April 2023

Abstract: Human activity recognition (HAR) is crucial to infer the activities of human beings, and to provide support in various aspects such as monitoring, alerting, and security. Distinct activities may possess similar movements that need to be further distinguished using contextual information. In this paper, we extract features for context-aware HAR using a convolutional neural network (CNN). Instead of a traditional CNN, a combined 3D-CNN, 2D-CNN, and 1D-CNN was designed to enhance the effectiveness of the feature extraction. Regarding the classification model, a weighted twin support vector machine (WTSVM) was used, which had advantages in reducing the computational cost in a high-dimensional environment compared to a traditional support vector machine. A performance evaluation showed that the proposed algorithm achieves an average training accuracy of 98.3% using 5-fold cross-validation. Ablation studies analyzed the contributions of the individual components of the 3D-CNN, the 2D-CNN, the 1D-CNN, the weighted samples of the SVM, and the twin strategy of solving two hyperplanes. The corresponding improvements in the average training accuracy of these five components were 6.27%, 4.13%, 2.40%, 2.29%, and 3.26%, respectively.

Keywords: ablation studies; context-awareness; convolutional neural network; feature extraction; human activity recognition; twin support vector machine

1. Introduction

Human activity recognition (HAR) has played an important role in various applications such as sport performance [1], life logging [2], anti-crime security [3], fall detection [4], health monitoring [5], and elderly care [6]. Typically, three types of sensors are utilized to measure the human activities, namely environment-based sensors, object-based sensors, wearable sensors, and video-based sensors [7]. Examples of environment-based sensors are radar, sound, and pressure detectors, as well as thermocouples, and barometers. Examples of object-based sensors are Wi-Fi and RFID. For wearable sensors, examples are the global positioning system, magnetometers, gyroscopes, and accelerometers. The forms of human activities vary from static (e.g., sitting and lying) to dynamic (e.g., walking and running) to transitive postures (transitions between consecutive activities). Since some human activities may share similar characteristics, contextual information is included to support the inferring of actual activities [8]. This is related to the concept of context-awareness, which is the ability to capture information from the external environment. Common contextual information includes the task, process, role, user, time, activity, identity, and location [9]. In this paper, HAR using video-based sensors is not considered in the performance comparison because the data type (as videos/images) is different from other types. Readers could refer to the works [10,11] for more details on HAR using video-based sensors.

Various machine learning algorithms have been proposed for context-aware HAR. Generally, there are five types of algorithms, namely the fuzzy logic-based [12,13], probabilistic-based [14,15], rule-based [16,17], distance-based [18,19], and optimization-based approaches [20,21].

To begin the discussion with fuzzy logic-based algorithms, a fuzzy rule-based inference system using fuzzy logic (FL) was proposed for the HAR of six activities: exercising, laying, sitting down, standing up, walking, and sleeping [12]. An experimental analysis of the system with a one day case study showed an accuracy of 97%. Another fuzzy rule-based inference system was presented for HAR with first person video [13]. The selected activities in the kitchen for the analysis were cleaning, washing dishes, and cooking. Biased classification was observed based on the evaluation metrics of recall (54.7%), precision (58.4%), and accuracy (70%).

Regarding the probabilistic-based algorithms, naive Bayes (NB) was chosen to take in the features of the fundamental DC component of a fast Fourier transform, as well as the value and variances of the magnitude, pitch, and roll for HAR [14]. Five human activities were included: walking, standing up, standing, laying down, and sitting. An accuracy of 89.5% was reported. A hidden Markov model (HMM) [15] was used to analyze 22 activities for HAR. The activities were: two-leg jump, one-leg jump, shuffle-right, shuffle-left, V-cut-right-right-first, V-cut-right-left-first, V-cut-left-left-first, run, spin-right-right-first spin-right-left-first, spin-left-right-first, spin-left-left-first, walk-downstairs, walk-upstairs, walk-curve-right, walk-curve-left, walk, stand-to-sit, sit-to-stand, stand, and sit. Owing to the complexity of the recognition of many activities, the model achieved an accuracy of 84.5%. A recent dissertation reported an extensive analysis of an innovative proposal using a motion-unit-based hidden Markov model [22]. The results confirmed that the approach (with a recognition rate over 90%) outperformed many existing works using benchmark datasets, including CSL-SHARE, ENABL3S, and UniMiB SHAR. It also outperformed another newly proposed method using various deep learning-based approaches, including convolutional neural networks, long short-term memory, and ResNet [23]. Attention is also drawn to a novel approach for automatic speech recognition and kinesiology to construct an HAR model that ensures expandability, generalizability, interpretability, and effective model training [24].

One of the rule-based HAR models was built using a random forest (RF) [16]. The model considered six activities, namely walking upstairs, walking downstairs, walking, standing, sitting, and laying. Researchers collected samples from 30 volunteers for the experimental analysis. The results had a sensitivity of 98%, a precision of 98.5%, an accuracy of 98%, and F1-score of 98%. Another approach was built using a decision tree (DT) [17]. Ten activities, namely walking, standing, sitting, eating, drinking (standing), drinking (sitting), sitting (standing), smoking (sitting), smoking (standing), and smoking (walking) were defined for analysis. For static activities, the model achieved an accuracy of 72%, whereas, for dynamic activities, an accuracy of 78% was observed.

K-nearest neighbor (KNN) is one of the most common distance-based algorithms. In [18], the authors enhanced the KNN algorithm with random projection to recognize 13 human activities: pushing a wheelchair, jumping, jogging, going downstairs, going upstairs, turning right, turning left, walking right (circle), walking left (circle), walking forward, lying down, sitting, and standing. The HAR model achieved an accuracy of 92.6%. Another work [19] applied two variants of KNN, namely fuzzy KNN and evidence theoretic

KNN for the HAR of twenty-nine activities using three benchmark datasets. The achieved accuracies were 77%, 93%, and 97% for datasets with 14, 10, and 5 activities, respectively.

For the optimization-based approach, a support vector machine (SVM) with a radial basis function was proposed [20]. Six human activities were considered, namely laying, standing, sitting, walking downstairs, walking upstairs, and walking. The experimental result showed that the model achieved an accuracy of 96.6%. In [21], an artificial neural network (ANN) model was built for the HAR of six human activities. The activities were laying, standing, sitting, walking, walking downstairs, and walking upstairs. An accuracy of 96.7% was observed.

A detailed pipeline was proposed for HAR research [25]. The key components involved are devices, sensors, software, data acquisition, segmentation, annotation, biosignal processing, feature extraction, feature study, activity modeling, training, recognition, evaluation, and application. Table 1 summarizes the key elements (sensors, feature extraction, method, context awareness, dataset, activities, cross-validation, and results) of the existing works [12–21]. The key research limitations of the existing works include the following:

- There are two major types of feature extraction. Major works [12–17,19–21] utilized the traditional feature extraction process. There has been less discussion (e.g., [18]) of automatic feature extraction using a deep learning algorithm, which may extract more representative features and eliminate the domain knowledge of all human activities;
- The methodology in the existing works [12–21] utilized traditional classification algorithms that may not work well with the nature of a high-dimensional feature space;
- Context awareness was omitted in the design and formulation of most of the existing works [12–16,18–21];
- Experimental analyses used limited benchmark datasets (and, thus, limited types of activities), with one dataset in most of the works [12–18,20,21] and three in [19]; and
- Cross-validation was omitted in most works [12–14,16–18,20,21]. It is important to fine-tune the hyperparameters and to evaluate the issue of model overfitting.

Table 1. Summary of the performance of existing works on HAR.

Work	Sensors	Feature Extraction	Method	Context Awareness	Dataset	Activities	CV	Results
[12]	Physiological, infrared debit, and state-change sensors; microphones	Raw sensor data	FL	No	One-day of data (simulated data)	Exercising, laying, sitting down, standing up, walking, and sleeping	No	Accuracy: 97%
[13]	Head-mounted camera	Motion coefficient	FL	No	Epic kitchens dataset (3088 samples) [26]	Cleaning, washing dishes, and cooking	No	Accuracy: 70% Precision: 58.4% Recall: 54.7%
[14]	Accelerometer	Value and variance of the magnitude, pitch, and roll; fundamental DC component of FFT	NB	No	16,5633 samples [27]	Walking, standing-up, standing, laying down, and sitting	No	Accuracy: 89.5%
[15]	Airborne microphone, electrogoniometer accelerometer, electromyography, and gyroscope	Built-in function using ASK2.0	HMM	No	CSL-SHARE dataset [28]	Two-leg jump, one-leg jump, shuffle-right, shuffle-left, V-cut-right-right-first, V-cut-right-left-first, V-cut-left-left-first, run, spin-right-right-first, spin-right-left-first, spin-left-right-first, spin-left-left-first, walk-downstairs, walk-upstairs, walk-curve-right, walk-curve-left, walk, stand-to-sit, sit-to-stand, stand, and sit	five-fold	Accuracy: 84.5%

Table 1. *Cont.*

Work	Sensors	Feature Extraction	Method	Context Awareness	Dataset	Activities	CV	Results
[16]	Gyroscope and accelerometer	Time –frequency domain analysis	RF	No	New dataset (30 volunteers)	Walking upstairs, walking downstairs, walking, standing, sitting, and laying	No	Accuracy: 98% F1-score: 98% Sensitivity: 98% Precision: 98.5%
[17]	Gyroscope and accelerometer	Absolute difference, correlation, integration, range, median, kurtosis, root-mean-square skewness, standard deviation, mean, maximum, and minimum	DT	Yes	3 month dataset (11 volunteers) [29]	Walk, stand, sit, eat, drink (standing), drink (sitting), sit (standing), smoke (sitting), smoke (standing), and smoke (walking)	No	Accuracy: 72% (static activities) Accuracy: 78% (dynamic activities)
[18]	Accelerometer, gyroscope	CNN	KNN with random projection	No	Wearable action recognition database [30,31]	Push wheelchair, jump, jog, go downstairs, go upstairs, turn right, turn left, walk right (circle), walk left (circle), walk forward, lie down, sit, and stand	k-fold (unspecified k)	Accuracy: 92.6%
[19]	Motion, temperature, phone usage, door, and pressure sensors	Weighted features from all sensors	Evidence theoretic KNN and fuzzy KNN	No	Kyoto1, Kyoto7, and Kasteren	Clean, cook, eat, phone call, wash hands, bed to toilet, prepare breakfast, groom, sleep, work at computer, work at dining room table, groom, prepare dinner, prepare lunch, watch tv, leave the house, the use toilet, take shower, obtain snack, obtain a drink, use washing machine, and wash dishes	LOO	Accuracy: 97% (Kyoto1) Accuracy: 77% (Kyoto7) Accuracy: 93% (Kasteren)
[20]	Accelerometer and gyroscope	Time–frequency domain analysis	SVM	No	10,299 samples [32]	Laying, standing, sitting, walking downstairs, walking upstairs, and walking	No	Accuracy: 96.6%
[21]	Accelerometer	Cyclic attribution technique	ANN	No	UCI-HAR (30 volunteers) [33]	Laying, standing, sitting, walking, walking downstairs, and walking upstairs	No	Accuracy: 96.7%

1.1. Research Contributions

To address the limitations of the research works, we propose a combined 3D-CNN, 2D-CNN, and 1D-CNN algorithm (3D-2D-1D-CNN) for feature extraction and a weighted twin support vector machine (WTSVM) for the HAR model. The contributions of the algorithm are explained as follows:

- The 3D-2D-1D-CNN algorithm leverages the ability of automatic feature extraction. An ablation study confirms that the 3D-CNN, 2D-CNN, and 1D-CNN achieve accuracy improvements of 6.27%, 4.13%, and 2.40%, respectively;
- The WTSVM takes the advantage of high-dimensional feature space and outperforms the twin SVM by 3.26% in terms of accuracy;
- Context awareness is incorporated to enhance the formulation of the HAR model, with an accuracy improvement of 2.4%; and
- Compared to existing works, our proposed algorithm enhances the accuracy by 0.1–40.1% with an increase of the total number of activities by 230–3100%.

1.2. Paper Organization

This paper is structured with five sections. Section 2 presents the details of the methodology of the 3D-2D-1D-CNN and WTSVM. The experimental results of the algorithm and its comparison are presented in Section 3. To investigate the contributions of the individual elements of the algorithm, ablation studies are carried out. A conclusion is drawn, with directions of future work, in Section 4.

2. Methodology

The methodology of the proposed HAR model comprises a feature extraction module using the 3D-2D-1D-CNN and a classification module using the WTSVM. In the following two subsections, the design and formulations are shared.

2.1. Feature Extraction Module Using the 3D-2D-1D-CNN

The feature extraction module is indispensable in machine learning. Traditional statistical [34,35] and time–frequency approaches [36,37] are not employed, since these approaches may not be effective in fully extracting representative features for a complex context-aware HAR problem with many human activities. Therefore, automatic feature extraction via CNN was chosen. Feature extractions via the 3D-CNN [38], 2D-CNN [39], and 1D-CNN [40] have been used in different applications with satisfactory performances. The features extracted using these algorithms may differ from each other. Combing the algorithms and, thus, merging more representative features is expected to further enhance the performance of the context-aware HAR model.

Figure 1 shows the ensemble architecture of the proposed 3D-2D-1D-CNN algorithm for feature extraction, which supports the classification model using the WTSVM. Three input layers are required to handle the inputs for the 3D-CNN, 2D-CNN, and 1D-CNN, as in a sub-model of the 3D-2D-1D-CNN algorithm. The workflows of the 3D-CNN, 2D-CNN, and 1D-CNN share similarities, with convolution operations, max pooling operations, and flatten operations.

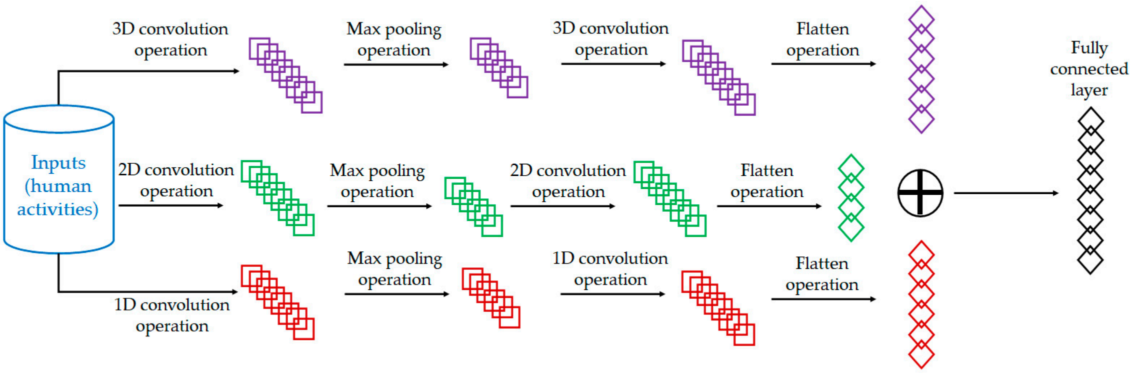

Figure 1. The architecture of the 3D-2D-1D-CNN algorithm for feature extraction.

The convolution operation acts on two samples to generate outputs representing the pattern change between the samples. The concept of pooling aims to reduce the data size, the training complexity of the model, and model overfitting. The maximum pooling operation takes the maximum values in the blocks to extract significant features. The flatten operation converts the data into a 1D matrix that is then passed into the fully connected layer. The features obtained in each algorithm will be merged into a fully connected layer. The design takes advantage of the effective coordination of training individual 3D-CNN, 2D-CNN, and 1D-CNN. It also facilitates the extraction of more representative features to enhance the performance of the HAR model. Ablation studies will be carried out to verify the contributions of the proposed ensemble architecture.

2.2. Classification Module Using a WTSVM

The basis of a weighted support vector machine (WSVM) for an imbalanced binary classification problem [41]—the hyperplane (or decision function)—is defined as

$$h(x) = w^T x + b = 0 \quad (1)$$

with weighted vector $w \in R^N$ and bias $b \in R$.

The constrained problem of a WSVM with a maximum margin hyperplane is given by:

$$\begin{gathered}\min_{w,b,\beta_-,\beta_+} \tfrac{1}{2}\|w\|^2 + p_-\beta_- + p_+\beta_+ \\ \text{s.t. } X_-w + c_-b \le c_- - \beta_- \\ X_+w + c_+b \le c_+ - \beta_+ \\ \beta_- \ge 0 \\ \beta_+ \ge 0\end{gathered} \tag{2}$$

where the regularization term $\frac{1}{2}\|w\|^2$ represents the maximum margin of the two parallel hyperplanes; p_- and p_+ are the penalty parameters to control the weights between terms for the negative class and positive class, respectively; β_- and β_+ are the slack variables for the negative class and positive class, respectively; X_- and X_+ are the training matrices for the negative class and positive class, respectively; and c_- and c_+ are the vectors for the negative class and positive class, respectively. In general, larger Lagrange multipliers may be assigned to some support vectors that help reduce the negative impact of imbalanced classification in an imbalanced dataset (details are provided in Section 3.1).

On the other hand, the basic formulations for the twin support vector machine (TSVM) are illustrated as follows. Different from WSVM, TSVM considers two non-parallel hyperplanes, $h_-(x)$ and $h_+(x)$:

$$h_-(x) = w_-^T x + b_- = 0 \tag{3}$$

$$h_+(x) = w_+^T x + b_+ = 0 \tag{4}$$

with the weighted vectors $w_- \in R^N$ and $w_+ \in R^N$ and the biases $b_- \in R$ and $b_+ \in R$. In these formulations, $h_-(x)$ is close to X_+ and far away from X_-, and $h_+(x)$ is close to X_- and far away from X_+.

The constrained problems of the TSVM are defined as:

$$\begin{gathered}\min_{w_-,b_-,\beta_+} \tfrac{1}{2}\|X_-w_- + c_-b_-\|^2 + p_1c_+^T\beta_+ \\ \text{s.t. } X_+w_- + c_+b_- \le c_+ - \beta_+ \\ \beta_+ \ge 0\end{gathered} \tag{5}$$

$$\begin{gathered}\min_{w_+,b_+,\beta_-} \tfrac{1}{2}\|X_+w_+ + c_+b_+\|^2 + p_2c_-^T\beta_- \\ \text{s.t. } X_-w_+ + c_-b_+ \le c_- - \beta_- \\ \beta_- \ge 0\end{gathered} \tag{6}$$

with the parameters $p_1 > 0$ and $p_2 > 0$.

Regarding our proposed WTSVM, it features (i) weights to adjust the level of the sensitivity of the hyperplane to respond to the imbalance ratio and (ii) majority points for the hyperplanes. To be practical, we formulated the WTSVM with a non-linear kernel function (K). The surfaces generated by the kernel function are given by

$$h_-(x) = K\left(x^T, X^T\right)w_- + b_- = 0 \tag{7}$$

$$h_+(x) = K\left(x^T, X^T\right)w_+ + b_+ = 0 \tag{8}$$

The optimization problems are defined as

$$\begin{gathered}\min_{w_-,b_-,\beta_+} \tfrac{1}{2}\left(\|w_-\|^2 + b_-^2\right) \\ +\tfrac{1}{2}p_1\left(\left(K(X_{1-},X)w_- + c_{1-}b_-\right)^T D_1(K(X_{1-},X)w_- + c_-b_-) + \beta_+^T\beta_+\right) \\ \text{s.t. } (K(X_+,X)w_- + c_+b_-) + \beta_+ \ge c_+ \\ \beta_+ \ge 0\end{gathered} \tag{9}$$

$$\begin{aligned} &\min_{w_+,b_+,\gamma_-} \tfrac{1}{2}\left(\|w_+\|^2 + b_+^2\right) \\ &+\tfrac{1}{2}p_2\left(\left(K(X_+,X)w_+ + c_+b_+\right)^T(K(X_+,X)w_+ + c_+b_+) + \beta_-^T D_2\beta_-\right) \\ &\text{s.t.} -(K(X_{2-},X)w_+ + c_{2-}b_+) + \beta_- \geq c_- \\ &\beta_- \geq 0 \end{aligned} \tag{10}$$

The Lagrange function of (9) is defined as

$$\begin{aligned} &L(w_-,b_-,\beta_+,\alpha) = \tfrac{1}{2}\left(\|w_-\|^2 + b_-^2\right) \\ &+\tfrac{1}{2}p_1\left(\left(K(X_{1-},X)w_- + c_{1-}b_-\right)^T(K(X_{1-},X)w_- + c_-b_-) + \beta_+^T D_1\beta_+\right) + \alpha^T(K(X_+,X)w_- + c_+b_- - \beta_+ + c_+) \end{aligned} \tag{11}$$

with the Lagrange multiplier $\alpha = \left(\alpha_1, \ldots, \alpha_{N_1}\right)^T$. Using Karush–Kuhn–Tucker conditions, we have

$$K(X_{1-},X)^T(K(X_{1-},X)w_- + c_-b_-) + p_1w_- + K(X_+,X)\alpha = 0 \tag{12}$$

$$c_-^T(K(X_{1-},X)w_- + c_-b_-) + p_1b_- + c_+^T\alpha = 0 \tag{13}$$

$$p_1\beta_+ - \alpha = 0 \tag{14}$$

$$(K(X_+,X)w_- + c_+b_-) + \beta_+ \geq c_+ \quad \beta_+ \geq 0 \tag{15}$$

$$\alpha^T((K(X_+,X)w_- + c_+)b_- - \beta_+ + c_+) = 0 \quad \alpha \geq 0 \tag{16}$$

The Lagrange dual problem of (9) becomes

$$\max_{\alpha} -\frac{1}{2}\alpha^T\left([K(X_+,X)c_+]\left[[K(X_{1-},X)c_-]^T[K(X_{1-},X)c_-] + p_1I\right]^{-1}[K(X_+,X)c_+]^T + \frac{1}{p_1}D_1^{-1}\right)\alpha + c_+^T\alpha \tag{17}$$

with $\alpha \geq 0$. Likewise, the Lagrange dual problem of (10) becomes

$$\max_{\alpha} -\frac{1}{2}\gamma^T\left([K(X_{2-},X)c_{2-}]\left[[K(X_+,X)c_+]^T[K(X_+,X)c_+] + p_2I\right]^{-1}[K(X_{2-},X)c_{2-}]^T + \frac{1}{2}D_2^{-1}\right)\gamma + c_{2-}^T\gamma \tag{18}$$

with the Lagrange multiplier $\gamma = \left(\gamma_1, \ldots, \gamma_{N_2}\right)^T$.

3. Performance Evaluation of the Proposed 3D-2D-1D-CNN-Based WTSVM for HAR

The proposed 3D-2D-1D-CNN-based WTSVM algorithm was evaluated using a benchmark dataset. This was followed by ablation studies to quantify the effectiveness of the individual components: 3D-CNN, 2D-CNN, 1D-CNN, weighted SVM, and twin SVM. We aimed to confirm that all components benefit the performance of the HAR model.

3.1. Dataset

To evaluate the performance of the proposed algorithm, a benchmark dataset, namely the ExtraSensory dataset, was chosen [42]. In this dataset, a mobile application was used to perform measurements of sixty volunteers for a one-minute recording with three twenty-second segments of various human activities. Motion-reactive sensors were used to collect data from the magnetometer, gyroscope, accelerometer, audio, compass, and location services. The dataset consisted of 308,306 labelled samples. Two types of labels were assigned to each record: (i) the primary activity (seven activities were included, namely bicycling, running, walking, standing and moving, standing in place, sitting, and lying down), and (ii) the secondary activity (a total of ninety-six specific contexts of various aspects, such as sleeping, eating, and cooking), which supplemented the primary activities. Each record could be linked with only one primary activity and multiple secondary activities. Figure 2 presents the number of samples of each of the primary activities. For activities defined as secondary activities, the number of samples is summarized in Table 2. As illustrative examples, Figure 3 shows the samples of the one-axis accelerometer readings with the phone-accelerometer (walking) and the watch-accelerometer, respectively.

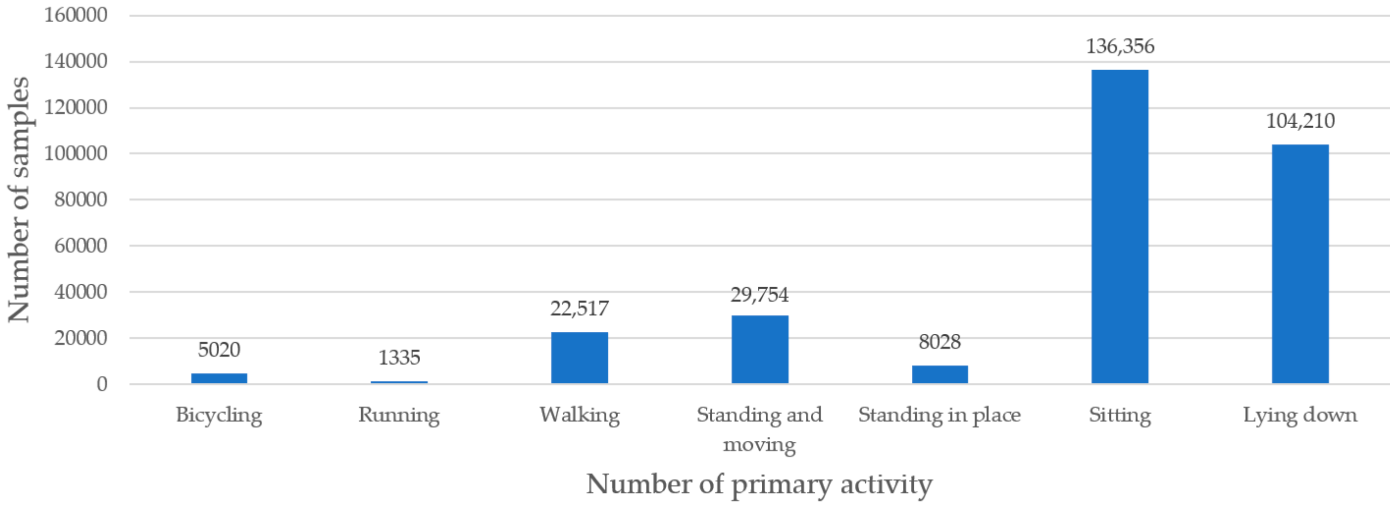

Figure 2. The number of samples of each of the primary activities.

Table 2. Summary of the number of samples of the secondary activities.

Name of Activity	No. of Samples	Name of Activity	No. of Samples	Name of Activity	No. of Samples	Name of Activity	No. of Samples
Phone on table	11,6425	At home	10,3889	Sleeping	83,055	Indoors	57,021
At school	43,221	Computer work	38,081	Talking	36,293	At work	29,574
Studying	26,277	With friends	24,737	Phone in pocket	24,226	Relaxing	21,223
Surfing the internet	19,416	Phone away from me	17,937	Eating	16,594	Phone in hand	16,308
Watching TV	13,311	Outside	11,967	Phone in bag	10,760	Listening to music with earphones	10,228
Written work	9083	Driving as driver	7975	With family	7975	With co-workers	6224
In class	6110	In a car	6083	Texting	5936	Listening to music without earphones	5589
Drinking non-alcohol	5544	In a meeting	5153	With a pet	5125	Listening to audio without earphones	4359
Reading a book	4223	Cooking	4029	Listening to audio with earphones	4029	Lab work	3848
Cleaning	3806	Grooming	3064	Exercising	2679	Toilet	2655
Driving as a passenger	2526	At a restaurant	2519	Playing videogames	2441	Laughing	2428
Dressing	2233	Shower bath	2087	Shopping	1841	On a bus	1794
Stretching	1667	At a party	1470	Drinking alcohol	1456	Washing dishes	1228
Smoking	1183	At the gym	1151	On a date	1086	Strolling	806
Going up the stairs	798	Going down the stairs	774	Singing	651	On a plane	630
Doing laundry	556	At a bar	551	At a concert	538	Manual labor	494
Playing phone games	403	On a train	344	Drawing	273	Elliptical machine	233
At the beach	230	At the pool	216	Elevator	200	Treadmill	164
Playing baseball	163	Lifting weights	162	Skateboarding	131	Yoga	128
Bathing	121	Dancing	115	Playing a musical instrument	114	Stationary bike	86
Motorbike	86	Transfer from bed to stand	73	Vacuuming	68	Transfer from stand to bed	63
Limping	62	Playing frisbee	54	At a sports event	52	Phone someone else using IT	41
Jumping	29	Phone strapped	27	Gardening	21	Ranking leaves	21
At sea	18	On a boat	18	Wheelchair	9	Whistling	5

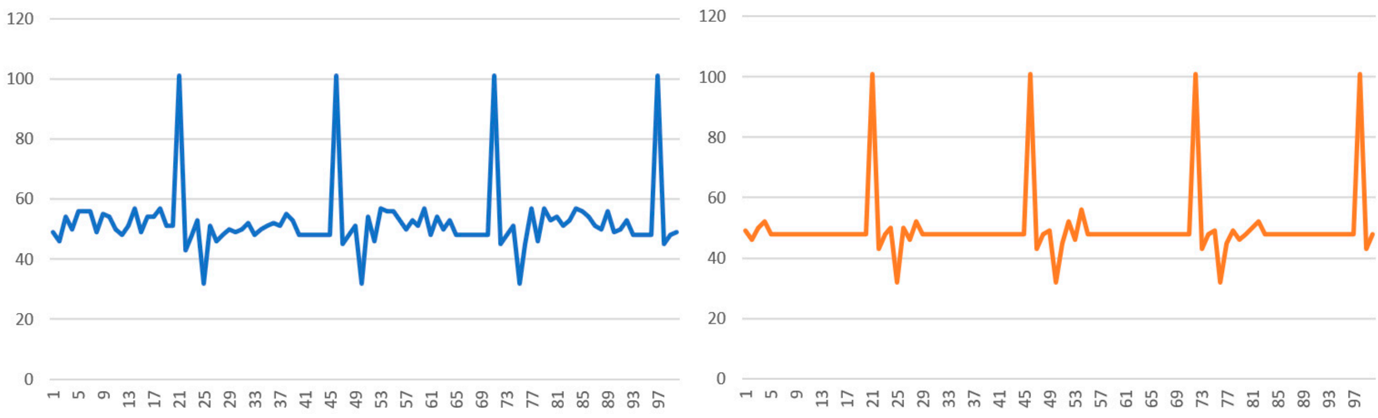

Figure 3. Samples of accelerometer readings. (**a**) Phone accelerometer (walking). (**b**) Watch accelerometer (walking).

3.2. Results

Owing to a few samples of the secondary activities of wheelchair (9) and whistling (5), a subject-wise K-fold cross-validation with K = 5 (instead of 10) was employed, which was also confirmed to be a common setting in the training and testing of machine learning models [43–45]. Four typical kernel functions, linear, radial basis, sigmoid, and polynomial, were chosen for the WTSVM model. Figure 4 shows the training accuracy and testing accuracy of the proposed 3D-2D-1D-CNN-based WTSVM using four kernel functions on the ExtraSensory dataset. The average training accuracy and average testing accuracy were 98.3% and 98.1%, respectively. There was only a small deviation between the average training and testing accuracy.

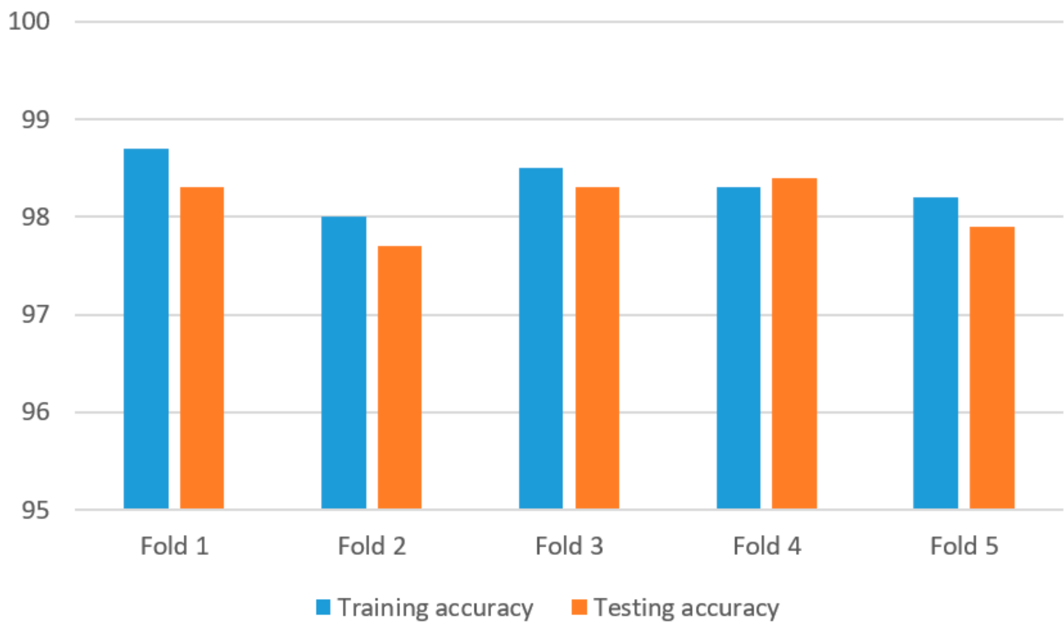

Figure 4. Accuracy of the 3D-2D-1D-CNN-based WTSVM on the ExtraSensory dataset.

To analyze the effectiveness of the 3D-2D-1D-CNN algorithm for feature extraction, an ablation study was carried out to study the contributions of the 3D-CNN, 2D-CNN, and 1D-CNN. The training accuracy, testing accuracy, precision, recall, and F1 score are summarized in Table 3. The key findings of the results are detailed as follows:

- The average training accuracy, average testing accuracy, average precision, average recall, and average F1 score were 98.3%, 98.1%, 98.4%, 98%, and 98.2%, respectively, for the 3D-2D-1D-CNN algorithm; 92.5%, 92.2%, 92.3%, 92.1%, and 92.2%, respectively, for the 2D-1D-CNN algorithm; 94.4%, 94.3%, 94.6%, 94.2%, and 94.3%, respectively, for the 3D-1D-CNN algorithm; and 96.0%, 95.9%, 96%, 95.8%, and 95.9%, respectively, for the 3D-2D-CNN algorithm. The results show that the average training accuracy was enhanced by 6.27%, 4.13%, and 2.40%, respectively;
- The ranking of the algorithms (from best to worst) based on the training accuracy and testing accuracy was 3D-2D-1D-CNN, 3D-2D-CNN, 3D-1D-CNN, and 2D-1D-CNN. This revealed the contributions of the individual components—the 3D-CNN, 2D-CNN, and 1D-CNN algorithms.

Table 3. Accuracy of the WTSVM using the 3D-2D-1D-CNN, 2D-1D-CNN, 3D-1D-CNN, and 3D-2D-CNN algorithms.

Method	Training Accuracy (%)/Testing Accuracy (%)/Precision (%)/Recall (%)/F1 Score (%)				
	Fold 1	Fold 2	Fold 3	Fold 4	Fold 5
3D-2D-1D-CNN	98.7/98.3/98.5/98.2/98.3	98.0/97.7/98.1/97.6/97.8	98.5/98.3/98.4/98.2/98.3	98.3/98.4/98.6/98.3/98.4	98.2/97.9/98.3/97.7/98.0
2D-1D-CNN	92.8/92.5/92.8/92.3/92.5	92.2/92.4/92.3/92.4/92.3	93.1/92.6/92.8/92.5/92.6	92.5/92.0/91.8/92.0/91.9	92.1/91.7/92.0/91.6/91.8
3D-1D-CNN	94.4/93.9/94.2/93.8/94.0	94.8/94.6/94.9/94.5/94.7	94.2/94.5/94.7/94.4/94.5	93.9/94.2/94.6/94.0/94.3	94.5/94.2/94.4/94.1/94.2
3D-2D-CNN	95.9/95.5/95.7/95.4/95.5	95.7/96.1/96.0/96.1/96.0	96.3/95.8/96.2/95.7/95.9	96.0/96.4/96.3/96.4/96.3	95.9/95.7/96.0/95.6/95.8

In addition, to analyze the effectiveness of the WTSVM, an ablation study was conducted to study the contributions of the weighting factors and the twin strategy. Table 4 summarizes the training accuracy and testing accuracy using the WTSVM, WSVM, and TSVM. The key observations are illustrated as follows:

Table 4. Accuracy of the 3D-2D-1D-CNN-based WTSVM, WSVM, and TSVM algorithms.

Method	Training Accuracy (%)/Testing Accuracy (%)				
	Fold 1	Fold 2	Fold 3	Fold 4	Fold 5
WTSVM	98.7/98.3/98.5/98.2/98.3	98.0/97.7/98.1/97.6/97.8	98.5/98.3/98.4/98.2/98.3	98.3/98.4/98.6/98.3/98.4	98.2/97.9/98.3/97.7/98.0
WSVM	95.5/95.8/96.0/95.7/95.8	94.9/94.6/94.4/94.7/94.5	95.0/94.5/94.7/94.4/94.5	95.4/95.8/96.0/95.7/95.8	95.0/94.6/94.9/94.5/94.7
TSVM	96.1/95.7/95.9/95.6/95.7	95.9/96.2/96.5/96.1/96.3	96.3/96.0/96.3/95.9/96.1	96.0/95.6/95.8/95.5/95.6	96.3/95.9/96.3/95.7/96.0

- The average training accuracy, average testing accuracy, average precision, average recall, and average F1 score were 98.3%, 98.1%, 98.4%, 98%, and 98.2%, respectively, for the WTSVM algorithm; 95.2%, 95.1%, 95.2%, 95.0%, and 95.1%, respectively, for the WSVM algorithm; and 96.1%, 95.9%, 96.2%, 95.8%, and 95.9%, respectively, for the TSVM algorithm. The enhancement of the average training accuracy by the WTSVM algorithm was 2.29% and 3.26%, respectively;
- The ranking of the algorithms (from best to worst) based on the training accuracy and testing accuracy was WTSVM, TSVM, and WSVM. This revealed the contributions of the individual components, the WTSVM, WSVM, and TSVM algorithms.

To evaluate the hypotheses, if the proposed approach outperformed all/some of the other approaches, a non-parametric Wilcoxon signed-rank test [46,47] was carried out. Table 5 summarizes the results of the five hypotheses. The p-values of all hypotheses were less than 0.05, suggesting that the proposed approach significantly outperformed the other approaches in the ablation studies.

Table 5. Results of hypothesis testing using the Wilcoxon signed-rank test.

Hypotheses	Results
H_0: 3D-2D-1D-CNN = 2D-1D-CNN; H_1: 3D-2D-1D-CNN > 2D-1D-CNN	Reject H_0
H_0: 3D-2D-1D-CNN = 3D-1D-CNN; H_1: 3D-2D-1D-CNN > 3D-1D-CNN	Reject H_0
H_0: 3D-2D-1D-CNN = 3D-2D-CNN; H_1: 3D-2D-1D-CNN > 3D-2D-CNN	Reject H_0
H_0: WTSVM = WSVM; H_1: WTSVM > WSVM	Reject H_0
H_0: WTSVM = TSVM; H_1: WTSVM > TSVM	Reject H_0

The performance of the proposed 3D-2D-1D-CNN-based WTSVM was compared with existing works [42,48–51] using the benchmark ExtraSensory dataset [42], and our work was evaluated using the benchmark datasets Kyoto1, Kyoto7, and Kasteren [30,31]. Table 6 summarizes the details of the comparisons. The key observations are summarized as follows:

- For ExtraSensory, our work enhanced the accuracy by 12.8–20.2% [42,48,51] and enhanced the F score by 17.2–86.0% [49,50];
- For Kyoto1, our work enhanced the accuracy by 0.918–2.89% [18,19];
- For Kyoto7, our work enhanced the accuracy by 8.89–13.1% [18,19];
- For Kasteren, our work enhanced the accuracy by 2.74–6.09% [18,19].

The results suggest that the proposed algorithm can manage varying scales of HAR problems with different numbers of activities. Regarding the computational complexity, the proposed algorithm requires more training time due to the complexity of the feature extraction process (3D-2D-1D-CNN). Nevertheless, this will not significantly affect the applicability of low-latency decisions in practice, because the classifier is based on WTSVM, a traditional machine learning classifier.

Table 6. Comparisons between existing works and our work.

Work	Methodology	Dataset	Number of Activities	Cross-Validation	Accuracy (%)
[42]	Early fusion	ExtraSensory [42]	25	5-fold	87
[48]	Random forest		15	10-fold	84
[49]	CNN with random forest		4	5-fold	52.8 (F score)
[50]	Deep graph CNN		25	N/A	83.8 (F score)
[51]	SVM		5	Single-split	81.6
Proposed	3D-2D-1D-CNN-based WTSVM		96	5-fold	98.1
[18]	Evidence-theoretic KNN and fuzzy KNN	Kyoto1 [30]	5	Leave-one-out	97
[52]	discriminative and generative SVM		5	Leave-one-out	98
Proposed	3D-2D-1D-CNN-based WTSVM		5	5-fold	98.9
[18]	Evidence-theoretic KNN and fuzzy KNN	Kyoto7 [30]	14	Leave-one-out	78
[52]	discriminative and generative SVM		14	Leave-one-out	81
Proposed	3D-2D-1D-CNN-based WTSVM		14	5-fold	88.2
[18]	Evidence-theoretic KNN and fuzzy KNN	Kasteren [31]	10	Leave-one-out	92
[52]	discriminative and generative SVM		10	Leave-one-out	95
Proposed	3D-2D-1D-CNN-based WTSVM		10	5-fold	97.6

4. Conclusions

It is desired for a context-aware HAR to be accurate and able to support the recognition of many activities. In this paper, we proposed 3D-CNN, 2D-CNN, and 1D-CNN algorithms (3D-2D-1D-CNN) for feature extraction and a weighted twin support vector machine (WTSVM) for an HAR model. A performance evaluation was carried out using four benchmark datasets. The proposed algorithm achieves an average training accuracy of 98.3% and an average testing accuracy of 98.1%. To investigate the contributions of the five individual components, namely the 3D-CNN, the 2D-CNN, the 1D-CNN, the weighted samples of the SVM, and the twin strategy of solving two hyperplanes, ablation studies were conducted. The results show the enhancement of the average training accuracy by 6.27%, 4.13%, 2.40%, 2.29%, and 3.26%, respectively, by the five individual components. In addition, we compared our work to 10 existing works. The comparison showed that our work enhanced the accuracy in four benchmark datasets.

The research team would like to suggest future research directions as follows: (i) generate more samples in the small classes, as in Table 1, using data generation algorithms such as the family of generative adversarial networks to reduce the issue of imbalanced recognition [53,54]; (ii) conduct online learning with new classes [55,56]; (iii) consider time-series image data [57,58] as inputs for the 3D-2D-1D-CNN algorithm; (iv) consider advanced feature extraction methods via high-level features [59], stacking, and feature space reduction [60,61]; and (v) investigate other non-training-based statistical methods of low-cost HAR algorithms [62,63].

Author Contributions: Formal analysis, K.T.C., B.B.G., M.T.-R., V.A., W.A. and I.F.Z.; investigation, K.T.C., B.B.G., M.T.-R., V.A., W.A. and I.F.Z.; methodology, K.T.C.; validation, K.T.C. and B.B.G.; visualization, K.T.C. and B.B.G.; writing—original draft, K.T.C., B.B.G., M.T.-R., V.A., W.A. and I.F.Z.; writing—review and editing, K.T.C., B.B.G., M.T.-R., V.A., W.A. and I.F.Z. All authors have read and agreed to the published version of the manuscript.

Funding: This research work was funded by Institutional Fund Projects under grant no. (IFPIP:1119-611-1443). The authors gratefully acknowledge technical and financial support provided by the Ministry of Education and King Abdulaziz University, DSR, Jeddah, Saudi Arabia.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

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

References

1. Hendry, D.; Chai, K.; Campbell, A.; Hopper, L.; O'Sullivan, P.; Straker, L. Development of a human activity recognition system for ballet tasks. *Sports Med.-Open* **2020**, *6*, 10. [CrossRef] [PubMed]
2. Jalal, A.; Quaid, M.A.K.; Tahir, S.B.U.D.; Kim, K. A study of accelerometer and gyroscope measurements in physical life-log activities detection systems. *Sensors* **2020**, *20*, 6670. [CrossRef] [PubMed]
3. Ullah, F.U.M.; Muhammad, K.; Haq, I.U.; Khan, N.; Heidari, A.A.; Baik, S.W.; de Albuquerque, V.H.C. AI-Assisted Edge Vision for Violence Detection in IoT-Based Industrial Surveillance Networks. *IEEE Trans. Ind. Inform.* **2022**, *18*, 5359–5370. [CrossRef]
4. Yadav, S.K.; Luthra, A.; Tiwari, K.; Pandey, H.M.; Akbar, S.A. ARFDNet: An efficient activity recognition & fall detection system using latent feature pooling. *Knowl.-Based Syst.* **2022**, *239*, 107948.
5. Bhavanasi, G.; Werthen-Brabants, L.; Dhaene, T.; Couckuyt, I. Patient activity recognition using radar sensors and machine learning. *Neural Comput. Appl.* **2022**, *34*, 16033–16048. [CrossRef]
6. Shu, X.; Yang, J.; Yan, R.; Song, Y. Expansion-squeeze-excitation fusion network for elderly activity recognition. *IEEE Trans. Circuits Syst. Video Technol.* **2022**, *32*, 5281–5292. [CrossRef]
7. Dang, L.M.; Min, K.; Wang, H.; Piran, M.J.; Lee, C.H.; Moon, H. Sensor-based and vision-based human activity recognition: A comprehensive survey. *Pattern Recognit.* **2020**, *108*, 107561. [CrossRef]
8. Miranda, L.; Viterbo, J.; Bernardini, F. A survey on the use of machine learning methods in context-aware middlewares for human activity recognition. *Artif. Intell. Rev.* **2022**, *55*, 3369–3400. [CrossRef]
9. Khowaja, S.A.; Yahya, B.N.; Lee, S.L. CAPHAR: Context-aware personalized human activity recognition using associative learning in smart environments. *Hum.-Centric Comput. Inf. Sci.* **2020**, *10*, 35. [CrossRef]

10. Tsai, J.K.; Hsu, C.C.; Wang, W.Y.; Huang, S.K. Deep learning-based real-time multiple-person action recognition system. *Sensors* **2020**, *20*, 4758. [CrossRef]
11. da Costa, V.G.T.; Zara, G.; Rota, P.; Oliveira-Santos, T.; Sebe, N.; Murino, V.; Ricci, E. Dual-head contrastive domain adaptation for video action recognition. In Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision, Waikoloa, HI, USA, 3–8 January 2022.
12. Medjahed, H.; Istrate, D.; Boudy, J.; Dorizzi, B. Human activities of daily living recognition using fuzzy logic for elderly home monitoring. In Proceedings of the 2009 IEEE International Conference on Fuzzy Systems, Jeju, Republic of Korea, 20–24 August 2009.
13. Schneider, B.; Banerjee, T. Bridging the Gap between Atomic and Complex Activities in First Person Video. In Proceedings of the 2021 IEEE International Conference on Fuzzy Systems, Luxembourg, 11–14 July 2021.
14. Maswadi, K.; Ghani, N.A.; Hamid, S.; Rasheed, M.B. Human activity classification using Decision Tree and Naive Bayes classifiers. *Multimed. Tools Appl.* **2021**, *80*, 21709–21726. [CrossRef]
15. Hartmann, Y.; Liu, H.; Schultz, T. Interactive and Interpretable Online Human Activity Recognition. In Proceedings of the 2022 IEEE International Conference on Pervasive Computing and Communications Workshops and other Affiliated Events, Pisa, Italy, 21–25 March 2022.
16. Radhika, V.; Prasad, C.R.; Chakradhar, A. Smartphone-based human activities recognition system using random forest algorithm. In Proceedings of the 2022 International Conference for Advancement in Technology, Goa, India, 21–22 January 2022.
17. Agac, S.; Shoaib, M.; Incel, O.D. Context-aware and dynamically adaptable activity recognition with smart watches: A case study on smoking. *Comput. Electr. Eng.* **2021**, *90*, 106949. [CrossRef]
18. Fahad, L.G.; Tahir, S.F. Activity recognition in a smart home using local feature weighting and variants of nearest-neighbors classifiers. *J. Ambient Intell. Humaniz. Comput.* **2021**, *12*, 2355–2364. [CrossRef] [PubMed]
19. Nguyen, B.; Coelho, Y.; Bastos, T.; Krishnan, S. Trends in human activity recognition with focus on machine learning and power requirements. *Mach. Learn. Appl.* **2021**, *5*, 100072. [CrossRef]
20. Muralidharan, K.; Ramesh, A.; Rithvik, G.; Prem, S.; Reghunaath, A.A.; Gopinath, M.P. 1D Convolution approach to human activity recognition using sensor data and comparison with machine learning algorithms. *Int. J. Cogn. Comput. Eng.* **2021**, *2*, 130–143.
21. Myo, W.W.; Wettayaprasit, W.; Aiyarak, P. A cyclic attribution technique feature selection method for human activity recognition. *Int. J. Intell. Syst. Appl.* **2019**, *11*, 25–32. [CrossRef]
22. Liu, H.; Schultz, I.T. Biosignal Processing and Activity Modeling for Multimodal Human Activity Recognition. Ph.D. Thesis, Universität Bremen, Bremen, Germany, 2021.
23. Mekruksavanich, S.; Jantawong, P.; Jitpattanakul, A. A Deep Learning-based Model for Human Activity Recognition using Biosensors embedded into a Smart Knee Bandage. *Procedia Comp. Sci.* **2022**, *214*, 621–627. [CrossRef]
24. Liu, H.; Hartmann, Y.; Schultz, T. Motion Units: Generalized sequence modeling of human activities for sensor-based activity recognition. In Proceedings of the 2021 29th European Signal Processing Conference, Dublin, Ireland, 23–27 August 2021.
25. Liu, H.; Hartmann, Y.; Schultz, T. A Practical Wearable Sensor-based Human Activity Recognition Research Pipeline. In Proceedings of the 15th International Conference on Health Informatics, Online, 9–11 February 2022.
26. Damen, D.; Doughty, H.; Farinella, G.M.; Fidler, S.; Furnari, A.; Kazakos, E.; Moltisanti, D.; Munro, J.; Perrett, T.; Price, W.; et al. Scaling egocentric vision: The epic-kitchens dataset. In Proceedings of the European Conference on Computer Vision, Munich, Germany, 8–14 September 2018.
27. Ugulino, W.; Cardador, D.; Vega, K.; Velloso, E.; Milidiú, R.; Fuks, H. Wearable computing: Accelerometers' data classification of body postures and movements. In Proceedings of the Advances in Artificial Intelligence-SBIA 2012: 21th Brazilian Symposium on Artificial Intelligence, Curitiba, Brazil, 20–25 October 2012.
28. Liu, H.; Hartmann, Y.; Schultz, T. CSL-SHARE: A multimodal wearable sensor-based human activity dataset. *Front. Comput. Sci.* **2021**, *3*, 759136. [CrossRef]
29. Shoaib, M.; Scholten, H.; Havinga, P.J.; Incel, O.D. A hierarchical lazy smoking detection algorithm using smartwatch sensors. In Proceedings of the 2016 IEEE 18th International Conference on e-Health Networking, Applications and Services, Munich, Germany, 14–17 September 2016.
30. Rashidi, P.; Cook, D.J.; Holder, L.B.; Schmitter-Edgecombe, M. Discovering activities to recognize and track in a smart environment. *IEEE Trans. Knowl. Data Eng.* **2010**, *23*, 527–539. [CrossRef]
31. Van Kasteren, T.; Noulas, A.; Englebienne, G.; Kröse, B. Accurate activity recognition in a home setting. In Proceedings of the 10th International Conference on Ubiquitous Computing, Seoul, Republic of Korea, 21–24 September 2008.
32. Xu, L.; Yang, W.; Cao, Y.; Li, Q. Human activity recognition based on random forests. In Proceedings of the 2017 13th International Conference on Natural Computation, Fuzzy Systems and Knowledge Discovery, Guilin, China, 29–31 July 2017.
33. Anguita, D.; Ghio, A.; Oneto, L.; Parra, X.; Reyes-Ortiz, J.L. A public domain dataset for human activity recognition using smartphones. In Proceedings of the ESANN 2013 Proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning, Bruges, Belgium, 24–26 April 2013.
34. Prathiba, S.B.; Raja, G.; Bashir, A.K.; AlZubi, A.A.; Gupta, B. SDN-assisted safety message dissemination framework for vehicular critical energy infrastructure. *IEEE Trans. Ind. Inf.* **2022**, *18*, 3510–3518. [CrossRef]

35. Deveci, M.; Pamucar, D.; Gokasar, I.; Köppen, M.; Gupta, B.B. Personal mobility in metaverse with autonomous vehicles using Q-rung orthopair fuzzy sets based OPA-RAFSI model. *IEEE Trans. Intell. Transport. Syst.* **2022**, 1–10. [CrossRef]
36. Rhif, M.; Ben Abbes, A.; Farah, I.R.; Martínez, B.; Sang, Y. Wavelet transform application for/in non-stationary time-series analysis: A review. *Appl. Sci.* **2019**, *9*, 1345. [CrossRef]
37. Elgendy, I.A.; Zhang, W.Z.; He, H.; Gupta, B.B.; Abd El-Latif, A.A. Joint computation offloading and task caching for multi-user and multi-task MEC systems: Reinforcement learning-based algorithms. *Wireless Netw.* **2021**, *27*, 2023–2038. [CrossRef]
38. Zhang, H.; Li, Y.; Jiang, Y.; Wang, P.; Shen, Q.; Shen, C. Hyperspectral classification based on lightweight 3-D-CNN with transfer learning. *IEEE Trans. Geosci. Remote Sens.* **2019**, *57*, 5813–5828. [CrossRef]
39. Chui, K.T.; Gupta, B.B.; Chi, H.R.; Arya, V.; Alhalabi, W.; Ruiz, M.T.; Shen, C.W. Transfer learning-based multi-scale denoising convolutional neural network for prostate cancer detection. *Cancers* **2022**, *14*, 3687. [CrossRef] [PubMed]
40. Hakim, M.; Omran, A.A.B.; Inayat-Hussain, J.I.; Ahmed, A.N.; Abdellatef, H.; Abdellatif, A.; Gheni, H.M. Bearing Fault Diagnosis Using Lightweight and Robust One-Dimensional Convolution Neural Network in the Frequency Domain. *Sensors* **2022**, *22*, 5793. [CrossRef]
41. Huang, Y.M.; Du, S.X. Weighted support vector machine for classification with uneven training class sizes. In Proceedings of the 2005 International Conference on Machine Learning and Cybernetics, Guangzhou, China, 18–21 August 2005.
42. Vaizman, Y.; Ellis, K.; Lanckriet, G. Recognizing detailed human context in the wild from smartphones and smartwatches. *IEEE Pervasive Comput.* **2017**, *16*, 62–74. [CrossRef]
43. Tembhurne, J.V.; Almin, M.M.; Diwan, T. Mc-DNN: Fake news detection using multi-channel deep neural networks. *Int. J. Semant. Web Inf. Syst.* **2022**, *18*, 1–20. [CrossRef]
44. Lv, L.; Wu, Z.; Zhang, L.; Gupta, B.B.; Tian, Z. An edge-AI based forecasting approach for improving smart microgrid efficiency. *IEEE Trans. Ind. Inf.* **2022**, *18*, 7946–7954. [CrossRef]
45. Chui, K.T.; Gupta, B.B.; Liu, R.W.; Zhang, X.; Vasant, P.; Thomas, J.J. Extended-range prediction model using NSGA-III optimized RNN-GRU-LSTM for driver stress and drowsiness. *Sensors* **2021**, *21*, 6412. [CrossRef]
46. Obulesu, O.; Kallam, S.; Dhiman, G.; Patan, R.; Kadiyala, R.; Raparthi, Y.; Kautish, S. Adaptive diagnosis of lung cancer by deep learning classification using wilcoxon gain and generator. *J. Health. Eng.* **2021**, *2021*, 5912051. [CrossRef] [PubMed]
47. Khosravi, K.; Pham, B.T.; Chapi, K.; Shirzadi, A.; Shahabi, H.; Revhaug, I.; Prakash, I.; Bui, D.T. A comparative assessment of decision trees algorithms for flash flood susceptibility modeling at Haraz watershed, northern Iran. *Sci. Total Environ.* **2018**, *627*, 744–755. [CrossRef]
48. Asim, Y.; Azam, M.A.; Ehatisham-ul-Haq, M.; Naeem, U.; Khalid, A. Context-aware human activity recognition (CAHAR) in-the-Wild using smartphone accelerometer. *IEEE Sensors J.* **2020**, *20*, 4361–4371. [CrossRef]
49. Cruciani, F.; Vafeiadis, A.; Nugent, C.; Cleland, I.; McCullagh, P.; Votis, K.; Giakoumis, D.; Tzovaras, D.; Chen, L.; Hamzaoui, R. Feature learning for human activity recognition using convolutional neural networks: A case study for inertial measurement unit and audio data. *CCF Trans. Pervasive Comput. Interact.* **2020**, *2*, 18–32. [CrossRef]
50. Mohamed, A.; Lejarza, F.; Cahail, S.; Claudel, C.; Thomaz, E. HAR-GCNN: Deep Graph CNNs for Human Activity Recognition from Highly Unlabeled Mobile Sensor Data. In Proceedings of the 2022 IEEE International Conference on Pervasive Computing and Communications Workshops and other Affiliated Events, Pisa, Italy, 21–25 March 2022.
51. Tarafdar, P.; Bose, I. Recognition of human activities for wellness management using a smartphone and a smartwatch: A boosting approach. *Decis. Support Syst.* **2021**, *140*, 113426. [CrossRef]
52. Fahad, L.G.; Rajarajan, M. Integration of discriminative and generative models for activity recognition in smart homes. *Appl. Soft Comput.* **2015**, *37*, 992–1001. [CrossRef]
53. Lu, Y.; Chen, D.; Olaniyi, E.; Huang, Y. Generative adversarial networks (GANs) for image augmentation in agriculture: A systematic review. *Comput. Electron. Agric.* **2022**, *200*, 107208. [CrossRef]
54. Chui, K.T.; Gupta, B.B.; Jhaveri, R.H.; Chi, H.R.; Arya, V.; Almomani, A.; Nauman, A. Multiround Transfer Learning and Modified Generative Adversarial Network for Lung Cancer Detection. *Int. J. Intell. Syst.* **2023**, *2023*, 6376275. [CrossRef]
55. Zhang, J.; Wang, T.; Ng, W.W.; Pedrycz, W. KNNENS: A k-nearest neighbor ensemble-based method for incremental learning under data stream with emerging new classes. *IEEE Trans. Neural Netw. Learn. Syst.* **2022**. [CrossRef]
56. Khosravi, K.; Golkarian, A.; Tiefenbacher, J.P. Using optimized deep learning to predict daily streamflow: A comparison to common machine learning algorithms. *Water Resour. Manag.* **2022**, *36*, 699–716. [CrossRef]
57. Li, C.; Zhao, H.; Lu, W.; Leng, X.; Wang, L.; Lin, X.; Pan, Y.; Jiang, W.; Jiang, J.; Sun, Y.; et al. DeepECG: Image-based electrocardiogram interpretation with deep convolutional neural networks. *Biomed. Signal Process. Control* **2021**, *69*, 102824. [CrossRef]
58. Makimoto, H.; Höckmann, M.; Lin, T.; Glöckner, D.; Gerguri, S.; Clasen, L.; Schmidt, J.; Assadi-Schmidt, A.; Bejinariu, A.; Müller, P.; et al. Performance of a convolutional neural network derived from an ECG database in recognizing myocardial infarction. *Sci. Rep.* **2020**, *10*, 8445. [CrossRef] [PubMed]
59. Hartmann, Y.; Liu, H.; Lahrberg, S.; Schultz, T. Interpretable High-level Features for Human Activity Recognition. In Proceedings of the 15th International Conference on Bio-inspired Systems and Signal Processing, Online, 9–11 February 2022.
60. Hartmann, Y.; Liu, H.; Schultz, T. Feature Space Reduction for Human Activity Recognition based on Multi-channel Biosignals. In Proceedings of the 14th International Conference on Bio-inspired Systems and Signal Processing, Online, 11–13 February 2021.

61. Hartmann, Y.; Liu, H.; Schultz, T. Feature Space Reduction for Multimodal Human Activity Recognition. In Proceedings of the 13th International Conference on Bio-inspired Systems and Signal Processing, Valletta, Malta, 24–26 February 2020.
62. Folgado, D.; Barandas, M.; Antunes, M.; Nunes, M.L.; Liu, H.; Hartmann, Y.; Schultz, T.; Gamboa, H. Tssearch: Time series subsequence search library. *SoftwareX* **2022**, *18*, 101049. [CrossRef]
63. Rodrigues, J.; Liu, H.; Folgado, D.; Belo, D.; Schultz, T.; Gamboa, H. Feature-Based Information Retrieval of Multimodal Biosignals with a Self-Similarity Matrix: Focus on Automatic Segmentation. *Biosensors* **2022**, *12*, 1182. [CrossRef] [PubMed]