

## Article

# Human Activity Recognition through Recurrent Neural Networks for Human–Robot Interaction in Agriculture

Athanasios Anagnostis <sup>1,2,3</sup>, Lefteris Benos <sup>1</sup>, Dimitrios Tsaopoulos <sup>1</sup>, Aristotelis Tagarakis <sup>1</sup>,  
Naoum Tsolakis <sup>1</sup> and Dionysis Bochtis <sup>1,2,\*</sup>

<sup>1</sup> Institute for Bio-Economy and Agri-Technology (IBO), Centre of Research and Technology-Hellas (CERTH), 6th km Charilaou-Thermi Rd, GR 57001 Thessaloniki, Greece; a.anagnostis@certh.gr (A.A.); e.benos@certh.gr (L.B.); d.tsaopoulos@certh.gr (D.T.); a.tagarakis@certh.gr (A.T.); n.tsolakis@certh.gr (N.T.)

<sup>2</sup> farmB Digital Agriculture, Doiraniis 17, GR54639 Thessaloniki, Greece

<sup>3</sup> Department of Computer Science & Telecommunications, University of Thessaly, GR35131 Lamia, Greece

\* Correspondence: d.bochtis@certh.gr

**Abstract:** The present study deals with human awareness, which is a very important aspect of human–robot interaction. This feature is particularly essential in agricultural environments, owing to the information-rich setup that they provide. The objective of this investigation was to recognize human activities associated with an envisioned synergistic task. In order to attain this goal, a data collection field experiment was designed that derived data from twenty healthy participants using five wearable sensors (embedded with tri-axial accelerometers, gyroscopes, and magnetometers) attached to them. The above task involved several sub-activities, which were carried out by agricultural workers in real field conditions, concerning load lifting and carrying. Subsequently, the obtained signals from on-body sensors were processed for noise-removal purposes and fed into a Long Short-Term Memory neural network, which is widely used in deep learning for feature recognition in time-dependent data sequences. The proposed methodology demonstrated considerable efficacy in predicting the defined sub-activities with an average accuracy of 85.6%. Moreover, the trained model properly classified the defined sub-activities in a range of 74.1–90.4% for precision and 71.0–96.9% for recall. It can be inferred that the combination of all sensors can achieve the highest accuracy in human activity recognition, as concluded from a comparative analysis for each sensor’s impact on the model’s performance. These results confirm the applicability of the proposed methodology for human awareness purposes in agricultural environments, while the dataset was made publicly available for future research.

**Keywords:** sensor fusion; accelerometer; gyroscope; magnetometer; machine learning; deep learning; long short-term memory networks; robotics; safe human-robot interaction; situation awareness



**Citation:** Anagnostis, A.; Benos, L.; Tsaopoulos, D.; Tagarakis, A.; Tsolakis, N.; Bochtis, D. Human Activity Recognition through Recurrent Neural Networks for Human–Robot Interaction in Agriculture. *Appl. Sci.* **2021**, *11*, 2188. <https://doi.org/10.3390/app11052188>

Received: 29 January 2021

Accepted: 24 February 2021

Published: 2 March 2021

**Publisher’s Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 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

Agriculture employs a significant number of workers all over the world, particularly in the developing countries. The adoption of modern technologies, including Information and Communication Technologies (ICT), such as Internet of Things (IoT), Artificial Intelligence (AI), Farm Management Information Systems (FMIS), wearable computers and robotics, has led to the so-called “agriculture 4.0” [1,2]. However, despite the plethora of technological advances, which have arguably ameliorated the farmers’ living standards to a substantial degree, their safety is frequently underestimated. Epidemiological studies have identified several safety and health issues associated with the agricultural occupations [3,4]. Focusing on the non-fatal health issues, work-related musculoskeletal disorders (MSDs) have proved to be the most ordinary ones [5,6].

The most common manual operations in agriculture include harvesting, weeding, digging, pruning, sorting, and load lifting and carrying, with the last task having gained relatively little attention. This does not conform to the gravity of the matter, as has been

reported recently in [7]. According to the literature, during load carriage, the energy expenditure increases as the load and carriage distance increase, while the possibility of getting injured increases when the worker carries the load incorrectly [8]. Repetitive load carriage, usually combined with lifting, has strongly been related to knee injuries and cumulative stresses on the lumbar region that can result in lower back pain, which have reached epidemic proportions among farmers [9]. Overall, the physical symptoms associated with load carriage can include aches, soreness, and feeling of fatigue that, in turn, can reduce noticeably farmer's performance. Another important aspect is that pickers, in tree and other perennial high value crops, spend considerable amount of time in carrying the filled crates outside the field, which can be a long distance. Alternatively, this task is shouldered by other farm workers, consequently increasing the costs for the farm owners. Remarkably, approximately one-third of the total working time can be spent in carrying crates, as stressed in [8].

Toward meeting the challenges of both preventing farmers from musculoskeletal injuries and assuring safety, the adoption of a fleet of low-cost lightweight robots with high-load carrying capacity for transporting the crates from the site of picking to the end of the farmyard can take place. Nevertheless, this is a multidisciplinary problem, as it entails issues coming from several scientific topics such as Human–Robot Interaction (HRI), IoTs, economics, Machine Learning (ML), clinical, safety, and ergonomics. As a consequence, each feature should be investigated separately, at a first stage, before integrated efforts are made for accomplishing a safe and economically viable solution. The present study focuses on a very important aspect of safe HRI, namely the activity recognition of the workers that can increase the situation awareness. This concept constitutes an emerging scientific field in HRI research to eliminate errors that usually appear in complex tasks. Most of the time, humans are situation aware and thus can reliably display competent performance. This skill is highly required also for autonomous vehicles, and ongoing research is carried out toward this direction [10,11]. If a task necessitates cooperation, the participants (human and robot in this case) should synchronize their actions. This represents a key challenge that should be addressed in HRI, as highlighted by Benos et al. [12], who examined the safety and ergonomics during human–robot synergy in agriculture.

Human Activity Recognition (HAR) has received extensive attention as a result of the progress of advanced technologies such as ML [13] and IoT (e.g., Inertial Measurement Units (IMUs) and high accuracy Global Positioning Systems (GPS)) [14]. In particular, HAR relied on sensors has been applied to numerous areas including rehabilitation [15], sports [16], healthcare [17], and security monitoring [18]. HAR based on sensor data is considered to be more reliable compared to vision-based techniques, since the latter are affected by visual disturbances, such as lighting variability, and need fixed site implementations [19]. In contrast, wearable sensor-based techniques are suitable for real-time implementation, as they are easy to be deployed, cost effective, and are not site dependent [20]. Among the most commonly utilized sensors, accelerometers, magnetometers, and gyroscopes are often used [21–23]. Smartphones, having multi-sensor systems, are also gaining attention [24,25]. In general, multi-sensor data fusion has been proved to be more reliable than using a single sensor, as possible information losses or the imprecision of one sensor can be compensated by the other sensors [24]. Remarkably, a promising technology is Digital Twin (DT), where multi-physics modeling can be integrated together with data analytics. As a result of its ability to combine data processing tools (from the digital world) as well as data acquisition (from the physical world), DT have been proposed as a means of better identifying scenarios of high-risk to optimize risk assessment and, thus, workers' safety [26,27].

Processes, such as data preprocessing, describing data collection, data fusion, outlier removal, and noise reduction, precede all other processes that are involved in data-driven methodologies. These steps are required to bring the data into an optimal state, since they are derived from sensors that potentially can produce irregularities. Depending on the nature of the problem and the ML algorithm selected to tackle it, the data need to be further

processed and transformed into the appropriate shape for a series of mathematical or logical operations. In general, the purpose of a ML algorithm is to be able to produce a model that will fit the data in the best possible way so as to predict unknown examples with the highest accuracy. In the case of HAR, the aim of the ML algorithm is to learn the characteristic features of the signals collected from the on-body sensors in order to be able to classify the correct activity for a particular timeframe. Afterwards, important feature vectors are extracted to minimize the classification errors and computation time [28]. Finally, the classification phase serves to map the selected features into a set of activities by exploiting ML techniques. Through implementing ML algorithms, models can be developed via iterative learning from the extracted features, up until they are able to optimally model a process. ML has extensively been implemented in agriculture, thus offering valuable solutions to several tasks such as crop, livestock, water, and soil management [29], to mention but a few. As far as the HAR is concerned, a plethora of ML models have been utilized, such as Hidden Markov Model [30], Support Vector Machine [31], K-Nearest Neighbor [32], Naive Bayes [33], Decision Tree [34], and Long Short-Term Memory (LSTM) [35]. Nonetheless, the literature regarding the use of ML for automated recognition, by using the data from wearable sensors collected throughout agricultural operations, is very limited. Indicative studies are those of Patil et al. [36] (use of accelerometers to detect digging, harvesting, and sowing) and Sharma et al. [37–39] (use of GPS, accelerometers, and microphone sensors to detect harvesting, weeding, bed making, and transplantation).

The aim of the present study was to properly identify human activities related to a particular task, which is lifting a crate and placing it onto a robot suitable for agricultural operations with the use of ML algorithms (LSTM) for sequential data classification. Since the agricultural environment is a dynamic ecosystem, which is susceptible to unforeseeable situations [40], other human sub-activities comprising this task were also investigated, including standing still as well as walking with and without the crate. Two common lightweight agricultural Unmanned Ground Vehicles (UGVs) were used: Husky and Thorvald robots. For the purpose of gathering the data, 20 healthy participants took part in outdoor experimental sessions by wearing five IMU sensors (embedded with tri-axial accelerometers, gyroscopes, and magnetometers) in different body positions. To the best of our knowledge, no similar study exists. This investigation, through providing the activity “signatures” of the workers, has the potential to increase human awareness during HRI, thus contributing toward establishing an optimal ecosystem in terms of both cost savings and safety. Finally, the present dataset is made publicly available [41] for the sake of future examination by other researchers.

## 2. Materials and Methods

### 2.1. Experimentation Setup

The experimental tests were carried out in a farm in the region of Volos, in central Greece. The study involved 20 participants (13 male, 7 female), whose average age, height, and weight were 30.95 years (SD  $\approx$  4.85), 1.75 m (SD  $\approx$  0.08), and 75.40 kg (SD  $\approx$  17.20), respectively, where SD corresponds to the standard deviation. The participants’ demographics are summarized in Table 1. To be eligible for inclusion in the present investigation, all subjects should have not had any history of surgeries or sustained any musculoskeletal injury during the last year that could influence their performance. All participants, prior to any experimental procedure, had to complete an informed consent form that was approved by the Institutional Ethical Committee.

**Table 1.** Participants' demographics.

Variable	Mean $\pm$ Standard Deviation
Age	30.95 years $\pm$ 4.85
Height	1.75 m $\pm$ 0.08
Mass	75.40 kg $\pm$ 17.20
Gender	13 males, 7 females

Following informed consent, each participant had to perform a specific activity. In particular, they have to walk an unobstructed distance of 3.5 m, lift a crate, and carry it to their point of departure, where they have to place it on an immovable agricultural robot. This task can be divided in six continuous sequential sub-activities, namely:

1. Standing still until the signal is given to start;
2. Walking a distance of 3.5 m without carrying any crate;
3. Bending down to approach the crate;
4. Lifting the crate from the ground to an upright position;
5. Walking back the distance of 3.5 m with carrying the crate;
6. Placing the crate onto the robot.

For the purpose of the present study, two UGVs were utilized (Husky and Thorvald), which are usually used in outdoor environments (Figure 1) [42,43]. The two available UGVs correspond to a deposit height of the crate equal to 40 cm (Husky) and 80 cm (Thorvald). Furthermore, the crate was either empty (tare weight equal to 1.5 kg) or full with weight plates with a total mass (crate and plates) approximately equal to 20% of each participant's mass, similarly to [44,45]. The mass of the available weight plates was 1 and 2.5 kg for the purpose of easily adjusting the required mass to be lifted and carried. An open plastic crate, commonly used in agriculture, was used with handles on both sides at 28 cm height above its base. The dimensions of the crate were 31  $\times$  53  $\times$  35 cm (height  $\times$  width  $\times$  depth). Consequently, each participant carried out four sub-cases:

1. Empty crate—Husky;
2. Crate full of the required weight—Husky;
3. Empty crate—Thorvald;
4. Crate full of the required weight—Thorvald.



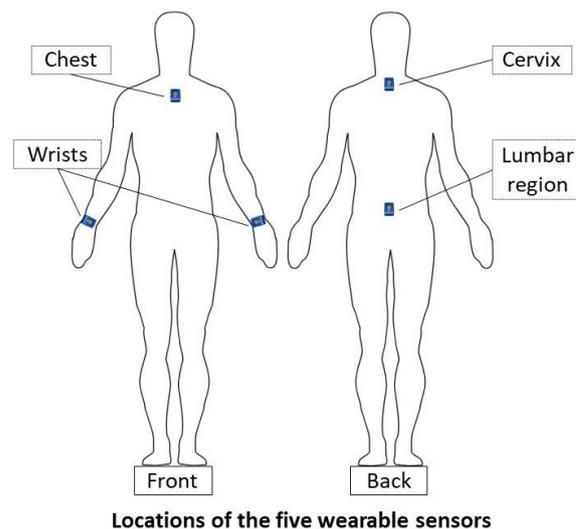
**Figure 1.** The available agricultural robots that were used in the present analysis, namely (a) Husky and (b) Thorvald robots.

Each sub-case was performed three times in a randomized order and at each participant's own pace, which stands for 12 efforts for each subject. Finally, all participants were instructed to carry out a five-minute warm-up in order to avoid possible injuries. The inclusion of different subjects, as per genre, age, weight, height, and loading heights on

robots, was targeted toward a large variability on the collected data, so the trained model was able to identify the activities conducted under a broad range of conditions.

## 2.2. Data Acquisition and Sensors

At the start of the day of experiments, the five VICON IMeasureU Blue Trident sensors were calibrated according to the manufacturer's directions [46]. These IMUs are small enough and lightweight (12 gr). Prior to the start of each effort, the sensors were attached to the chest (breastbone), cervix (approximately T1 vertebra), lumbar region (approximately L4), right and left wrist, as can be depicted in Figure 2. IMUs were attached via special Velcro straps at the two wrists (provided by the manufacturer), while the remaining three sensors were attached via double-sided tape similarly to studies such as [47]. Each IMU encompasses a tri-axial accelerometer, a tri-axial gyroscope, and a tri-axial magnetometer. The specifications of the IMUs are summarized in Table 2 according to [48]. These kinds of sensors have been used in several recent studies, including [49–52]. The sampling frequency that was used throughout the experimental sessions was 50 Hz, which is considered to be adequate for such kind of investigations, similarly to experimental investigations such as [21,24,52].



**Figure 2.** Locations of the five wearable sensors on the human body during experiments.

**Table 2.** The specifications of the VICON IMeasureU Blue Trident sensors, as mentioned in [48].

Sensor	Measured Magnitude	Range Units	Sensitivity	Axes
Accelerometer	Acceleration	$\pm 16$ g	16 bit	3 axes
Gyroscope	Angular velocity	$\pm 2000^\circ/\text{sec}$	16 bit	3 axes
Magnetometer	Direction/position	$\pm 4900$ $\mu\text{T}$	16 bit	3 axes

The Capture.U software (provided by VICON) [53] was used to synchronize the sensors and capture the data, while the latter were saved directly to the sensors for further processing. Since Capture.U is available only for iOS devices, an Apple iPad mini (64 GB) [54] was utilized for the present investigation.

## 2.3. Signal Preprocessing

### 2.3.1. Distinguishing the Sub-Activities

The Capture.U software in conjunction with the iPad offers the choice of simultaneously recording the experimental session at hand. This feature was particularly useful for distinguishing the sub-activities and finding out the critical instant, where the transition

between them took place similarly to [55]. Each sequence starts with the subject standing still (labeled as sub-activity “0”) and is used as a two-fold baseline; (a) for establishing a distinctive and realistic “idle” activity and (b) for allowing the timely synchronization of the sensors before starting the sequence. The rest of the sub-activities are described next:

- Walking without the crate sub-activity (labeled as “1”) begins immediately after one of the feet first leaves the ground (beginning of stance phase of gait [56]);
- The third sub-activity (labeled as “2”) begins when the participant starts bending their trunk, kneeling, or simultaneously performs them both, which are usually referred to as stoop, squat, and semi-squat techniques, respectively, in the relative literature [57];
- The fourth sub-activity (labeled as “3”) begins when the participant starts lifting the crate from the ground [58];
- The fifth sub-activity (labeled as “4”) begins when the participant starts the stance phase of gait similarly to the above description, however, by carrying the crate this time;
- The sixth sub-activity (labeled as “5”) begins when the participant starts bending their trunk, kneeling, or simultaneously performs them both, as described above, while it ends when the entire surface of the crate is placed onto the UGV.

Obviously, as the required tasks are continuous in nature, the beginning of one task corresponds to the end of the other. Since the participants performed the tasks in their own way and at their own pace, following the above well-defined criteria was of major importance in order to assure the reliability of the results. In this fashion, it should be highlighted that in case at least one of the sensors was not synchronized with the others (this was recognized during processing the dataset), the corresponding measurements of the remaining four sensors were also discarded. The aforementioned labeling of the sub-activities was used for the rest of the signal pre-processing, as will be elaborated below.

### 2.3.2. Outlier and NaN Handling

Under normal conditions, data collection from sensors in real-life applications can be problematic. Throughout experimentation, hardware failures and malfunctions can occur, resulting in gaps or irregular values in the dataset. During experimental sessions, sensors could randomly stop collecting data, thus creating gaps in the dataset. These gaps were identified in the early stages, and the entire experiment was discarded completely. On the other hand, in some cases, the sensors would record all data properly. However, during the pre-processing phase, irregular values, i.e., outliers, could appear in the dataset. These outliers usually differ from most values by orders of magnitude, and they do not represent the physical behavior of the subject. In the present study, there was a limited number of outliers, which were all removed manually from the dataset.

### 2.3.3. Noise Reduction

Signals or one-dimensional data usually contain unwanted or unknown components that can be the result of the capturing, transmitting, or storing device. In this investigation, the focus was on removing the noise from the collected data and not identifying its root. Noise removal techniques serve so as to remove irregular fluctuations by running filters (mathematical operations) throughout the entire signal and replace the “noisy” values with “smooth” ones. These methods enable the ML algorithms to better learn the trends and fluctuations instead of random value variations. In this analysis, the median filter [59] was used as a noise removal method. This kind of filter is a nonlinear one within which each output sample is calculated as the median value of input samples under the selected window. This corresponds to a result after the sorting of the input values. Furthermore, the median filtering of signals includes a horizontal window having an odd number of taps. After examining various values, eleven taps were utilized. In addition, no isolated extreme values (such as a possible large-valued sample as a result of impulse noise) appeared in the filtering phase, since outliers were removed manually beforehand. The median filter’s effect on the signals is demonstrated in Figure 3.

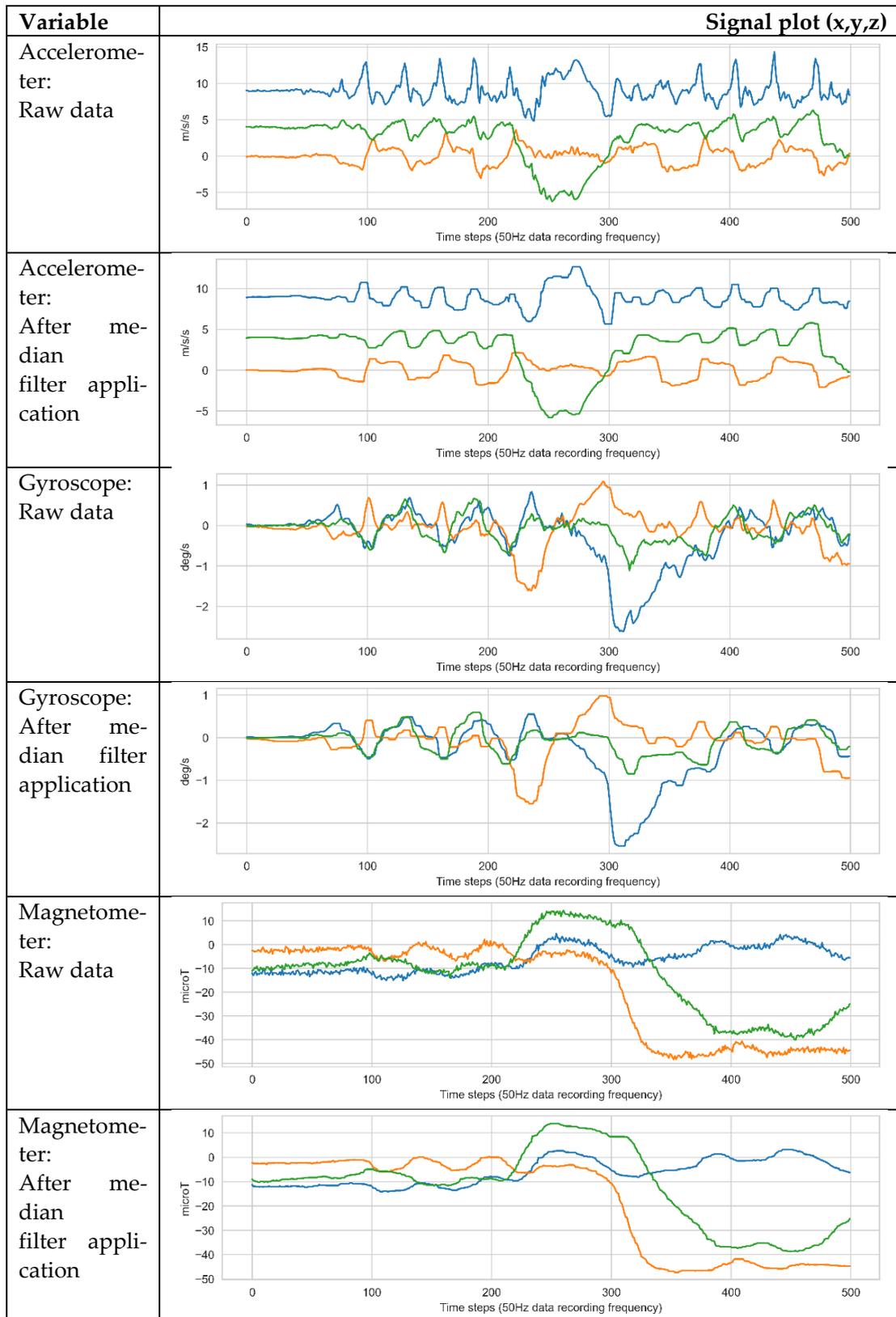
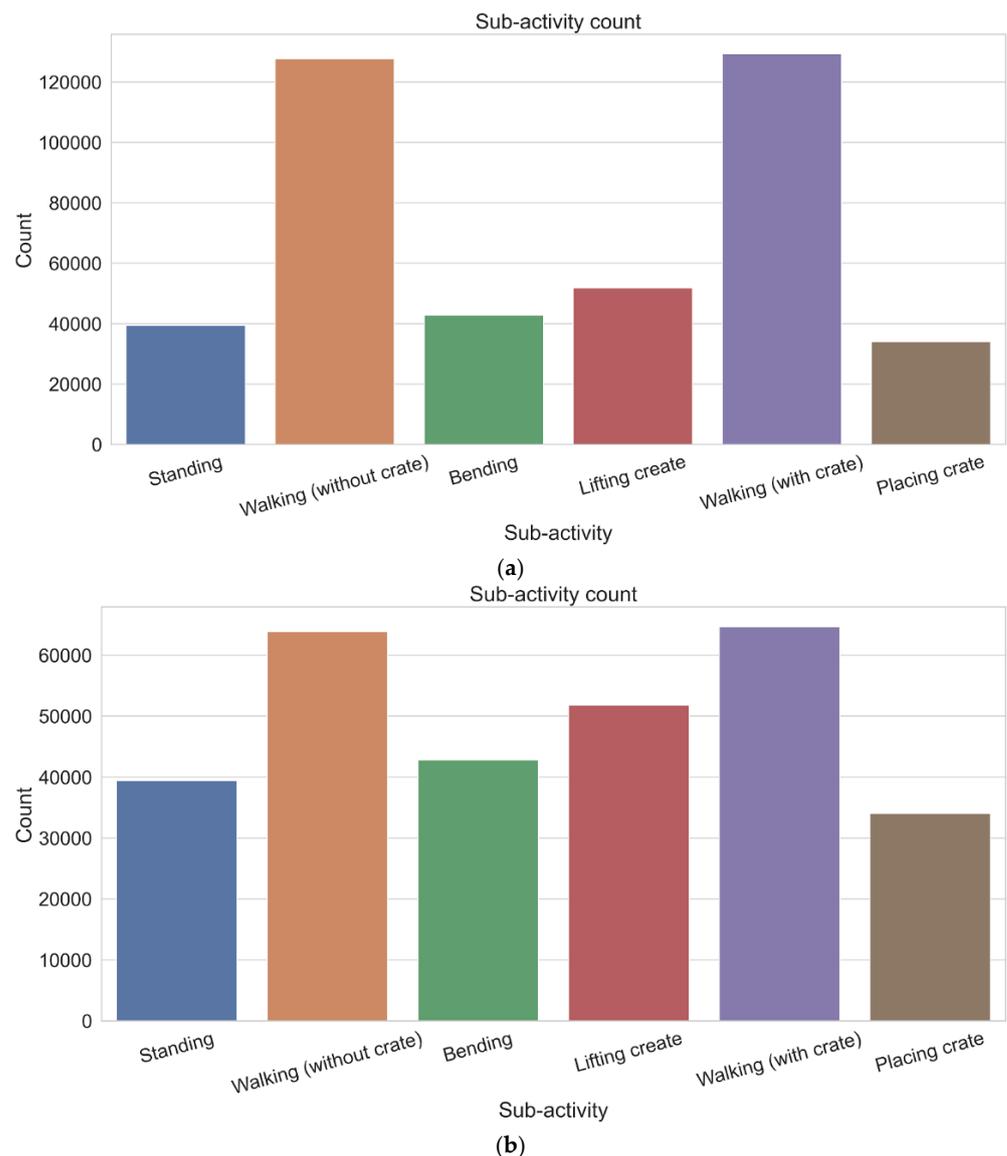


Figure 3. Collected raw signals and processed signals after median filter application.

#### 2.3.4. Activity Count and Class Imbalance

In the present study, the experimental sessions, which were designed for the participants, included six sub-activities, as described in Section 2.1. As expected, these sub-activities did not need the same time to be executed. More specifically, bending to approach the crate, lifting it from the ground to an upright position, or placing it on a loading surface were very short activities. On the contrary, walking with and without the crate lasted approximately triple time as compared to the other sub-activities, as shown in Figure 4a, as the crates and the robots were initially at 3.5 m with each other.



**Figure 4.** Histograms illustrating (a) counts of each sub-activity on the original dataset and (b) the result after the under sampling.

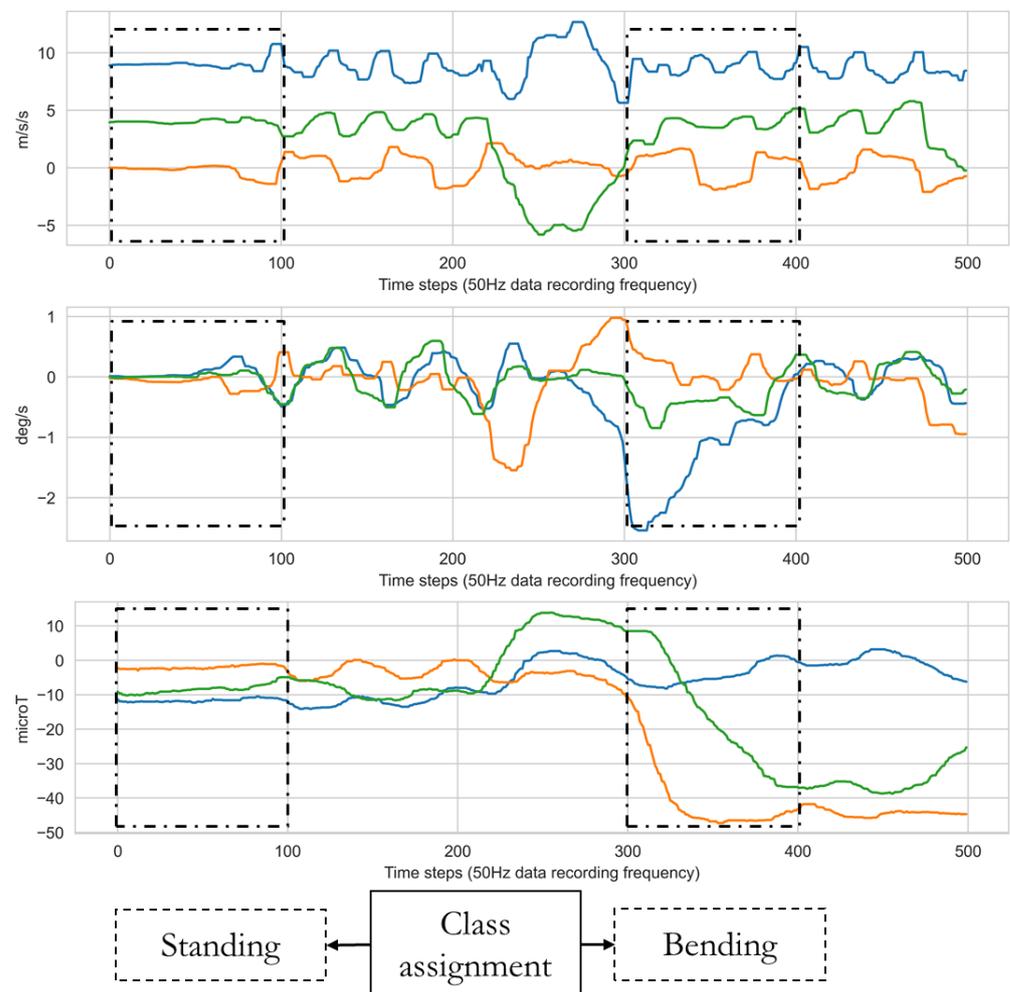
The observed imbalance, existing in the training set, may become troublesome. In fact, class imbalance can create problems pertaining to the ML algorithms' performance, particularly in cases where classes are identified that can be wrongly interpreted to each other owing to commonalities. For the purpose of reducing the existence of the above classes, with the least effect associated with the remaining information, an under-sampling technique was implemented. The under sampling was performed via taking out every other entry on the dataset in case the sub-activity matched one of the two most populated

classes. The classes do not include an equal number of instances. Nonetheless, there exists a better balance between them, as can be gleaned from Figure 4b.

## 2.4. Feature Engineering

### 2.4.1. Temporal Window Definition

As a temporal window, we define the time that is needed to identify an activity by the sensors' data. Depending on the addressed problem, this window can be set as small as 1 s [60], or as large i.e., >6 s [61]. The effect of the temporal window has been thoroughly investigated for a multitude of HAR problems; however, a similar approach defines the window at 2.56 sec [31]. For the present analysis, the temporal window was set to 2 s after extensive investigation of values that ranged from 0.5 to 5 s. For each temporal window, a class is assigned, representing the sub-activity that the subject was conducting for the particular time. An indicative schematic on the temporal window and class assignment is presented in Figure 5.

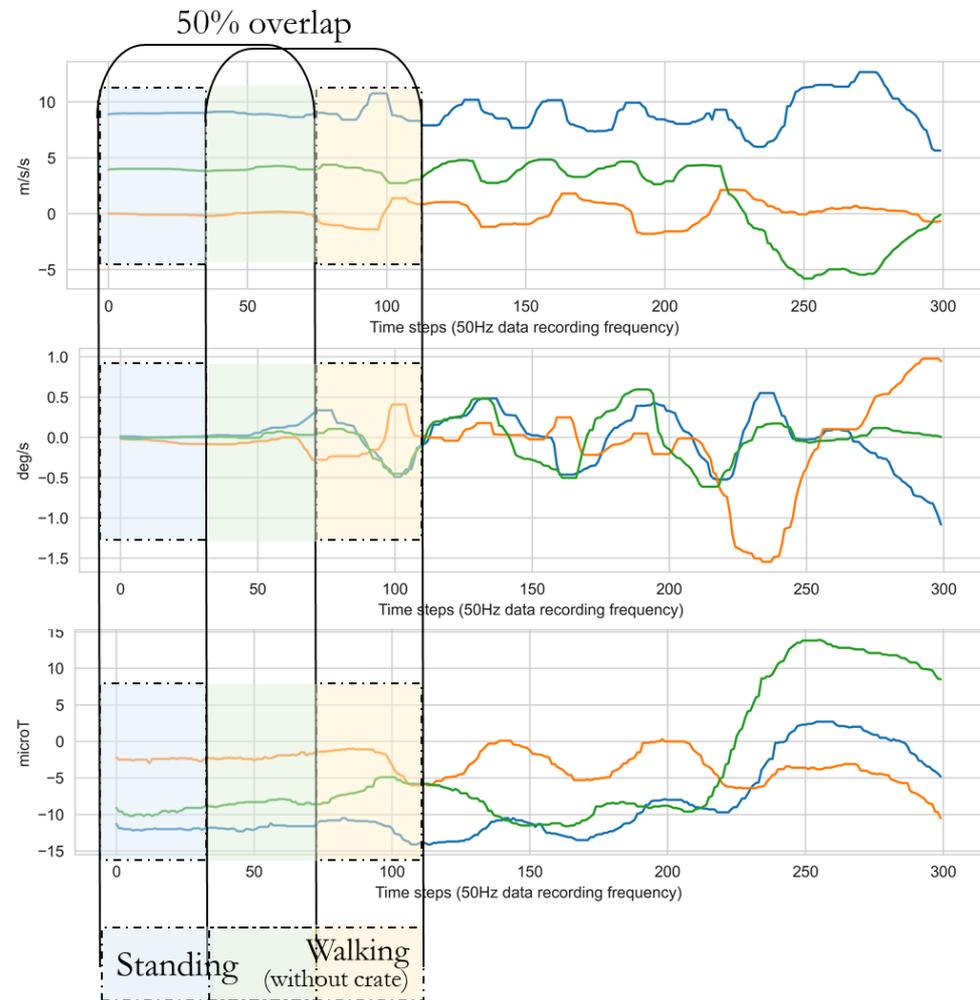


**Figure 5.** Temporal windows on signals and class assignment.

### 2.4.2. Overlap

Class assignment, based on temporal windows, can identify some activities accurately; however, there can be activities that fall between temporal windows and are not represented in either. With the intention of solving this issue, an overlapping of the temporal windows was conducted automatically for all signals. Consequently, every next temporal window does not start at the end of the previous one, but exactly in its middle in order to achieve 50% overlap. This technique offers double benefit: one being that it minimizes the chances

to miss an activity due to falling between windows, and second because it increases the training examples for the ML algorithm to train to. A representative schematic of the overlapping technique is presented in Figure 6.



**Figure 6.** Temporal window overlapping.

#### 2.4.3. Categorical Variables

By definition, the sub-activities that were assigned as classes to each temporal window are categorical variables, since they are descriptive qualitative variables. Each sub-activity's description was set as a class, and for each class, a number was assigned to simplify the data collection process. ML models work with categorical variables, but with the intention of including them into the calculations; they need to be transformed to numerical representations. Human intuition points toward the use of a simple integer representation as a numerical value to an abstract categorical variable. Nevertheless, integer numbers contain order and, therefore, they imply order and operations to each activity, i.e., walking with the crate is two times bending, which makes no sense. This issue can be solved by transforming the integers into one-hot vectors, which are of the same magnitude, i.e., one, and their length is the number of the sub-activities. These vectors contain zero values, except for one digit that is different for each activity. By using this technique, all numerical representations have the same magnitude, have no order between them, and can be used in the calculations of the ML algorithms. The sub-activity description, its assigned value, and the resulting vector are presented in Table 3.

**Table 3.** Sub-activity description, assigned value, and vectorial transformation.

Sub-Activity	Assigned Value	One-Hot Vector
Standing	0	(1,0,0,0,0)
Walking (without crate)	1	(0,1,0,0,0)
Bending	2	(0,0,1,0,0)
Lifting crate	3	(0,0,0,1,0)
Walking (with crate)	4	(0,0,0,0,1)
Placing crate	5	(0,0,0,0,1)

#### 2.4.4. Train/Test Split

The dataset was split into a training portion containing the examples, which will be used for training the model, and a testing portion containing the examples that will be used to evaluate the model's performance and robustness. The testing portion was completely removed from the dataset prior to any operation that would result in having the training set leak information and compromise the validity of the model's predictions. Due to the amount of data obtained from the experimentation phase, an 80/20 split was selected for the training/test datasets. The split was conducted on the subject level in order to evaluate the performance of the trained model on the specific characteristics of an unknown subject's movements. Thus, four subjects were randomly selected for testing purposes, i.e., to have the trained model predict their activities as recorded. The data from the remaining sixteen subjects are to be used for training the ML algorithm.

#### 2.4.5. Normalization

Normalization is the process where all features of a dataset are scaled into a common range. In the present study, the StandardScaler from Python's Sklearn library was utilised [62] which calculated the standard score as:

$$z = (x - u) / SD \quad (1)$$

where  $u$  is the mean of the training samples and  $SD$  is the standard deviation of the training samples.

Normalization is applied only to the training dataset so as to prevent information leaking toward the test dataset, which needs to be completely unknown during the training process. Better optimization during training can also be achieved via normalization, as it appropriately speeds up the convergence of the non-convex cost function to the global minimum.

### 2.5. Machine Learning Algorithm (LSTM)

LSTM networks constitute a type of neural network architecture being built upon a recurrent manner, via introducing memory cells, as well as the in-between connections, with the intention of constructing a graph directed in a sequence. In a general sense, recurrent neural networks process the sequences by employing these memory cells in a dissimilar way as compared to simple artificial neural networks. Although they are designed for handling problems with a sequential nature, recurrent neural networks frequently comfort the problem pertaining to vanishing gradients or being unable to "memorize" many sequential data. However, the characteristic cell structures found in LSTMs, also called gates, render the network capable of varying the retained information [63], while they can regulate which part of information is going to be either discarded or stored in the long memory. This leads to the much desired optimization of the memorizing process. Problems with dynamic sequential behavior have been proven to be suitable for such kinds of problems. HAR can fall under this premise, because all activities are time-dependent sequences, which makes LSTM a suitable algorithm for the problem the present study tackles.

## 2.6. Performance Metrics

In this subsection, the utilized performance metrics are briefly described. Generally, this type of metrics is employed with the objective of offering a common measure concerning the trained classifier's performance against the unknown examples originated from the testing set. The result of this prediction, as compared with the actual class label assigned to each activity, can acquire one of the following values:

- True Positive (TP) or True Negative (TN), in case that it is classified correctly;
- False Positive (FP) or False Negative (FN), in case that it is misclassified.

Subsequently, the aforementioned values are implemented as a means to compute the performance metrics, commonly appeared in classification problems [64]. In Table 4, the performance metrics, which were used in the present investigation for the sake of appraising the classifier's performance, are summarized in conjunction with a concise description of them and their mathematical relationship.

**Table 4.** Performance metrics used in the present study.

Name	Description	Formula
Accuracy	Ratio of correctly predicted observation to the total observations (preferred in balanced datasets)	$(TP+TN)/(TP+FP+FN+TN)$
Precision	Ratio of correctly predicted positive observations to the total predicted positive observations	$TP/(TP+FP)$
Recall	Ratio of correctly predicted positive observations to all observations in actual class	$TP/(TP+FN)$
F1 score	The weighted average of Precision and Recall (preferred in unbalanced datasets)	$(2 \times \text{Recall} \times \text{Precision})/(\text{Recall} + \text{Precision})$

Finally, for the purpose of assessing the performance of the present algorithm with respect to the given data, a loss function is utilized, which is also known as an objective or cost function. Since the present study deals with a multiclass classification problem, the categorical cross-entropy was adopted, which calculates the loss among the probability one-hot vectors of the real and the predicted class [65]. The mathematical formula for the cross-entropy loss that was used is given by:

$$\text{Cross-entropy} = -\sum p(x)\log q(x). \quad (2)$$

In the above equation,  $p(x)$  is the probability vector for the real class, and  $q(x)$  is the probability vector for the predicted class. This function minimizes the loss relative to how good the predictions are, and it increases the loss in a steep manner for bad predictions (reaching up to infinity when the prediction is completely wrong).

## 2.7. Proposed Machine Learning Pipeline

The complete ML pipeline of the proposed methodology is summarized in this section. The pipeline starts with the signal preprocessing, which includes the loading of the data, the fusion of variables and axes from all sensors, the removal of NaNs (i.e., Not a Number) and outliers, the removal of noise, and the balancing of the classes. The next step is the feature engineering and data transformation that is needed so that the data will be in proper condition for the training. This includes the temporal window definition, the overlap of temporal windows, the class assignment and labeling, the splitting of the training and test datasets, and the normalization of the training dataset. The following process is the ML model training with the use of an LSTM algorithm architecture, which learns each sub-activity's features the training dataset. Finally, validation of the model is done with a 10-fold cross-validation, where the training and test dataset shuffle after each training is over and the performance metrics are calculated. With the completion of the 10-fold cross-validation, the performance metrics are averaged and presented. A schematic of the pipeline is presented in Figure 7.

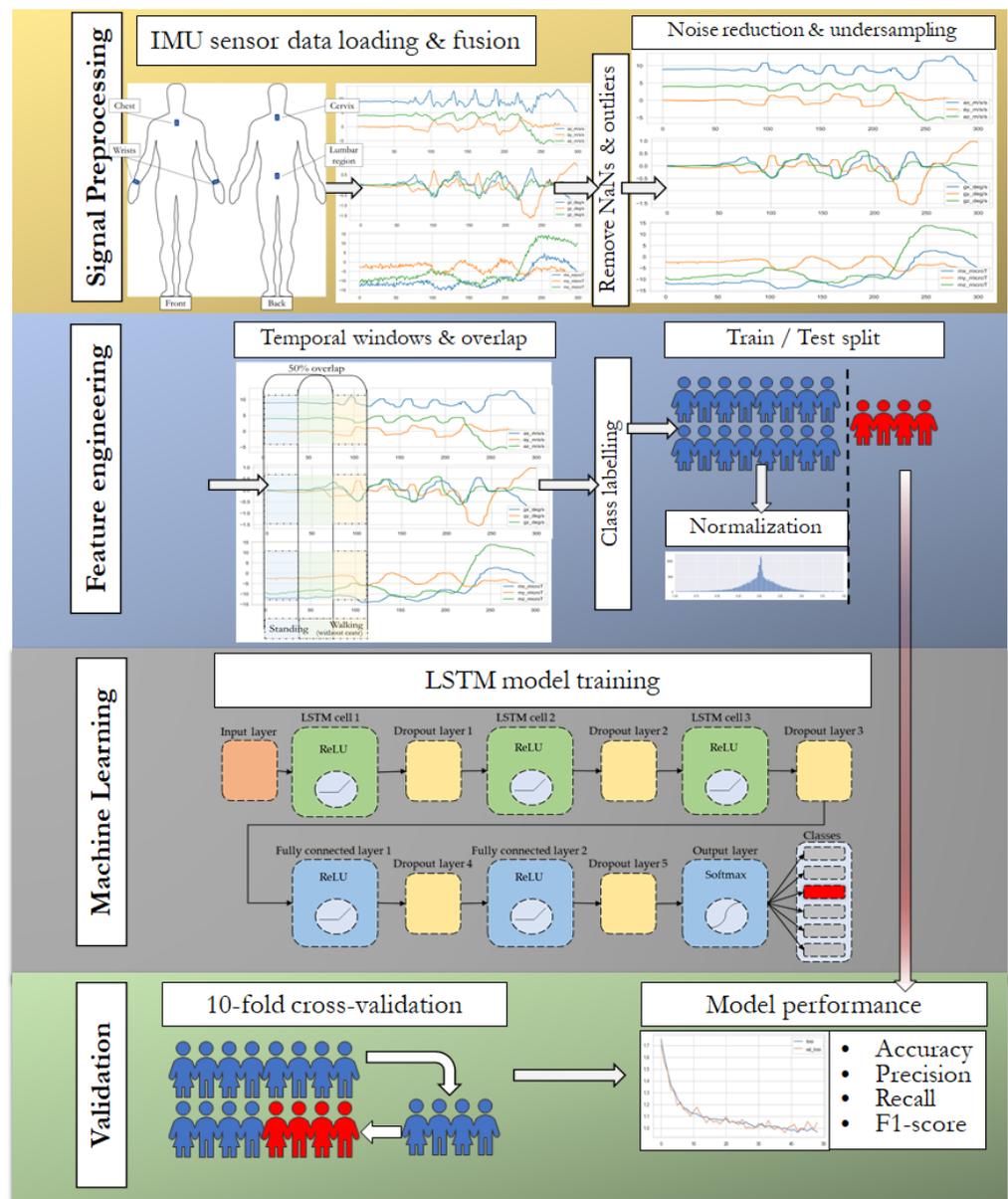


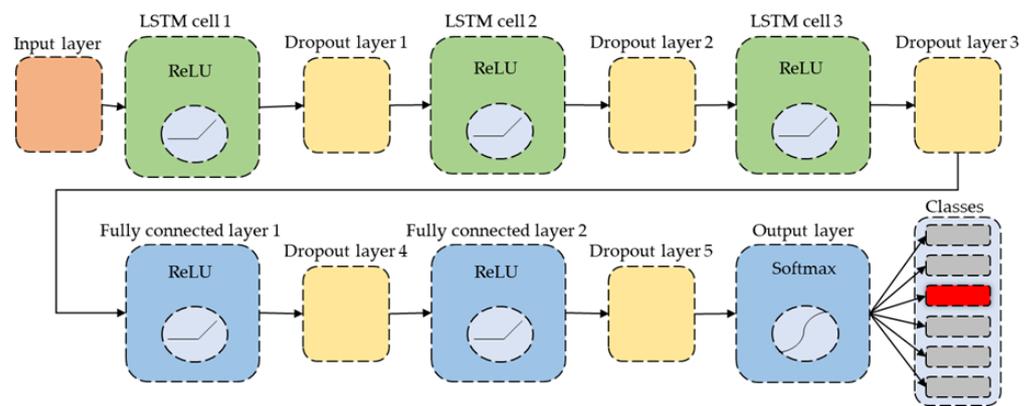
Figure 7. Schematic of the Machine Learning (ML) pipeline proposed in the present study.

### 3. Results

#### 3.1. LSTM Architecture

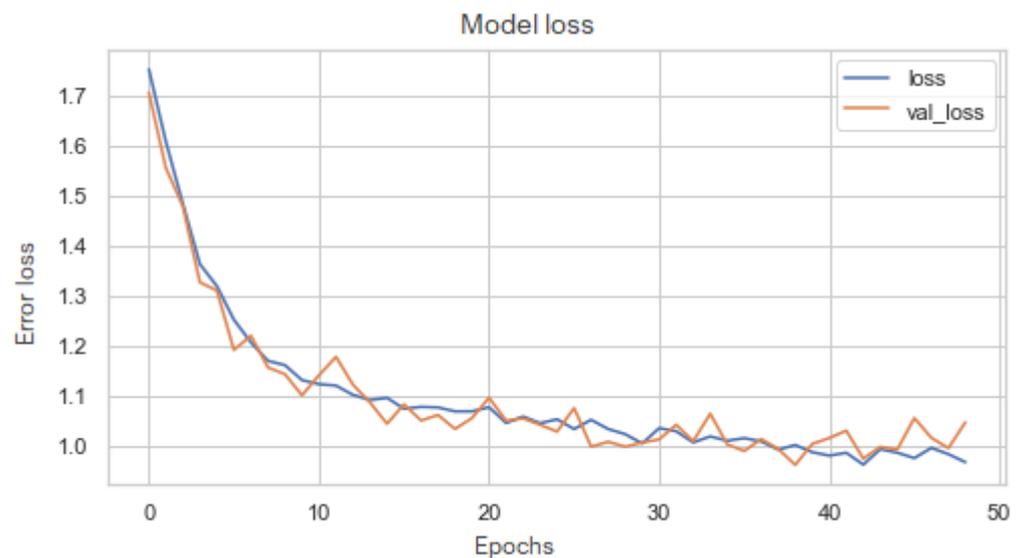
Extensive investigation and tryouts were conducted with the aim of identifying the optimal specifications and hyper-parameter values for the LSTM architecture. The selected architecture is illustrated in Figure 8.

The three LSTM cells are constructed by 10 memory units each. The first and second fully connected layers have 10 nodes. All the aforementioned layers use Rectified Linear Unit (ReLU) activation [66], and each one is followed by a dropout layer with 50% drop rate. The output layer has six nodes, one for each class and uses the softmax activation [67]. The total number of trainable parameters sums up to 12,701.



**Figure 8.** Schematic representation of the proposed Long Short-Term Memory (LSTM) architecture.

Training of the model ranged from 25 to 50 epochs. A typical plot of the training and validation loss decrease over the epochs is shown in Figure 9.



**Figure 9.** Training and validation plot loss during a complete training.

### 3.2. Confusion Matrix

Based on the results of the performance metrics, the model's performance can be visualized in confusion matrices. In particular, the confusion matrix is a table that displays the aforementioned values in such a way that one can easily view the number of properly classified examples, as well as false positives and false negatives. In this analysis, a multiclass classification, the confusion matrix is of size  $6 \times 6$ , where six is the number of activities that are predicted, as seen in Table 5.

### 3.3. Classification Report

The classification report displays the prediction, recall, and F1-score for each class. Individual performance metrics for each class is a useful tool to comprehend a model's weaknesses and strengths. The classes were engineered in the preprocessing phase to be generally balanced. Nevertheless, both the macro (not weighted) average and the weighted average of all metrics are calculated. Since the accuracy is calculated considering all predictions, there is only one value, and it describes the general performance of the trained model. The classification report is shown in Table 6.

**Table 5.** The confusion matrix for the present model's performance.

Confusion Matrix		Predicted Classes					
		Standing	Walking (Without Crate)	Bending	Lifting Crate	Walking (With Crate)	Placing Crate
True classes	Standing	135	5	0	0	1	7
	Walking (without crate)	2	227	5	0	0	0
	Bending	1	16	101	9	1	0
	Lifting crate	1	1	30	165	11	0
	Walking (with crate)	1	1	1	11	215	0
	Placing crate	12	0	0	0	31	106

**Table 6.** Classification report for the proposed model's performance.

Activity	Precision	Recall	F1-Score
Standing	0.887	0.919	0.902
Walking (without crate)	0.904	0.969	0.937
Bending	0.741	0.790	0.763
Lifting crate	0.888	0.788	0.833
Walking (with crate)	0.834	0.897	0.864
Placing crate	0.860	0.710	0.777
<b>Accuracy</b>			0.856
<b>Macro average</b>	0.853	0.843	0.845
<b>Weighted average</b>	0.860	0.856	0.854

Overall, the "Walking without crate" sub-activity presents the higher predictions by achieving 0.904 for precision, 0.969 for recall, and 0.937 for F1-score. In contrast, the "Bending" sub-activity achieves the lowest values for precision (0.741) and F1-score (0.763), while the "Placing crate" one achieves the lowest recall (0.710). The trained model's total performance is measured by the accuracy metric, which demonstrates a total of 85.6% for all the defined activities.

### 3.4. Feature Selection

Investigation regarding the effect each variable has on the performance of the model was also conducted. The accelerometer, gyroscope, and magnetometer data were used both individually and combined, in order to compare with the result, the model achieved when all variables were utilized. Next, the accuracy of each approach was measured along with its error, and the results are presented in Table 7.

**Table 7.** Feature selection for the model's variables and their respective accuracies.

Accelerometer (x,y,z)	Gyroscope (x,y,z)	Magnetometer (x,y,z)	Accuracy ( $\pm$ Error)
x			82.708% ( $\pm$ 1.468)
	x		82.987% ( $\pm$ 1.380)
		x	73.412% ( $\pm$ 2.180)
x	x		83.619% ( $\pm$ 0.877)
x		x	81.200% ( $\pm$ 1.261)
	x	x	85.532% ( $\pm$ 0.604)
x	x	x	85.677% ( $\pm$ 0.972)

As can be deduced from Table 7, the combination of all sensors performed better as compared to the case where each sensor was used individually. This was an expected result that has been highlighted by several relative studies, such as [24,68]. On the other hand, considering the usage of a single sensor, gyroscopes appear to slightly outweigh accelerom-

eters, demonstrating an accuracy of 82.987% and 82.708%, respectively. Concerning the magnetometer, it was observed to have the poorest performance, while its supplemental usage, as part of a case considering only two types of sensors, is suggested only in combination with a gyroscope, leading to approximately 3.07% increase of the accuracy. In contrast, the synergy of accelerometers and gyroscopes resulted in approximately 1.10% and 0.76% increase of the accuracy as compared to the purely usage of accelerometers and gyroscopes, respectively.

#### 4. Discussion and Main Conclusions

HAR is of major importance in the design process of agricultural collaborative robotic systems, as they should be able to operate in dynamic and crowded farm environments, where almost nothing is structured. In addition, these collaborative systems do not use isolated cells as occurs with conventional industrial robots. Toward optimizing the required activities, robots are working in the same working region concurrently with their “co-workers”, namely humans. “Cobots”, as these robots are usually referred to [69], can carry out either the same task or distinct tasks. The present envisioned application focuses on the latter scenario, where the robot can follow the workers while harvesting and, subsequently, place the crate onto the robot. Afterwards, the robot can safely transfer the full crates outside the field. Aside from the aim to provide safety and time saving, this cooperation can contribute to the prevention of the fatigue of agricultural workers, because the arduous task of carrying the crates for a long distance is performed by robots based on human-aware planning. Apart from the HRI, the results of this study are also applicable to conventional in-field operations such as lifting crates and loading to platforms for transferring to storage.

The activity recognition of workers is closely related to an essential feature of HRI, which is usually mentioned as “social-aware robot navigation” [70,71]. While autonomous navigation is restricted to obstacles’ avoidance and reaching the target destination [72–74], the social navigation, apart from that, takes into consideration other factors associated with human naturalness, comfort, and sociability [75,76]. More specifically, naturalness is related to navigation in paths such as those for humans via adjusting the robot’s speed and its distance from farmers. Comfort offers also the feeling of safety, whereas sociability has to do with abstracting decisions pertaining to robot’s movements by considering ethical and regional notions [71]. In a nutshell, HAR within agricultural human–robot ecosystems has a great potential to assure a sociable acceptable safe motion of robots and provide a free space to farmers to perform their activities unaffected by the simultaneous existence of robots, while the latter can approach them when it is required [12,77,78].

The present study focuses solely on HAR. To this end, data originated from 20 healthy participants, carrying out a particular task, were gathered by five wearable IMUs. This task included walking an unobstructed distance, lifting a crate (either empty or with a total mass of 20% of each participant’s body mass), and carrying it to the point of departure, where they have to place it onto an immovable UGV (either a Husky or a Thorvald). By carefully distinguishing the sub-activities comprising the above task, the obtained signals were properly preprocessed for the purpose of using them in the learning phase (training of the model) and the testing phase (evaluating the model’s performance and robustness) of an ML process (LSTM).

Overall, the problem of properly classifying stationary activities was challenging. The “Bending”, “Lifting”, and “Placing” activities were initially misclassified by a large margin. However, noise removal and normalization increased the overall performance of the trained model significantly. One of the factors that improved the performance of the model was the width of the temporal window which, when it varied more than one second, the performance of the overall model would decrease substantially. The LSTM architecture has provided the appropriate tools for the model to be able to learn the features of the activity signals. Early experimentation with artificial neural networks (ANN) and one-dimensional convolutional neural networks (CNN) has resulted in low performance on the trained model. Nevertheless, further investigation on more elaborate architectures

utilizing the benefits of multiple methods, such as CNN-LSTM [77] or convolutional LSTM [78] networks might be worth conducting. However, being a characteristic of data-driven approaches, the volume and variability of data play a significant role in a model's performance. That being stated, this study has shown that by obtaining data from 20 subjects, equipped with five IMU sensors each, performing a few recordings and fine-tuning a state-of-the-art LSTM network can help train a robust model, which can properly classify all activities with an accuracy of larger than 76%. Toward increasing the volume and variability of data (and, thus, the overall accuracy), a study with a sample that consisted of more participants covering a wider range of ages, physical strength, and anthropometric characteristics exists in the immediate plans of the authors. Moreover, with the intention of providing real-world data, these experimental tests are planning to be performed in a real agricultural environment by workers at their own pace, according to the complex conditions that they may face.

Additionally, it can be concluded that the gyroscope and the accelerometer can be both used independently for recognizing the specific sub-activities, which are commonly performed in agricultural environments. However, their synergetic contribution can somewhat increase the overall performance. In contrast, the use of a magnetometer alone cannot lead to equally reliable results and should only be considered for supplementary use. The best performance was presented when the data from all sensors were fused. Furthermore, for the sub-activity of walking without the crate, the present methodology indicated the higher precision. On the contrary, as anticipated, the sub-activity presenting the smallest precision was that of bending down in order to approach the crate, since it can be executed in several ways, depending on each participant. For example, it was observed that most of the time, participants could solely bend their trunks (stooping) or kneel without bending their trunk enough, or simultaneously stoop and kneel to catch the crate. This resulted from the instruction of participants to carry out the task in their own way. This is justified from our intention to increase the variability of the dataset for capturing, as widely as possible, most of the different manners in which someone can perform the desired task.

Obviously, assuring a fluid and safe HRI in agriculture involves a plethora of different issues. However, each issue must be addressed separately, at a preliminary stage, before a viable solution is proposed. This study demonstrates the framework for both conducting direct field measurements and applying a ML approach to accurately identify the activities of workers automatically, by analytically presenting the applied methodology at each phase. Finally, the examined dataset is made publicly available, thus assuring research transparency while allowing for experimental reuse and lowering the barriers for meta-studies.

**Author Contributions:** Conceptualization, L.B., A.A., D.T., N.T., D.B.; methodology, L.B., A.A., A.T.; software, A.A.; validation, A.A.; writing—original draft preparation, L.B., A.A.; writing—review and editing, L.B., A.A., D.T., A.T., N.T., D.B.; visualization, L.B., A.A.; supervision, D.T., D.B.; project administration, D.B. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Institutional Review Board Statement:** The study was conducted according to the guidelines of the Declaration of Helsinki, and approved by the Ethical Committee under the identification code 1660 on 3 June 2020.

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

**Data Availability Statement:** The dataset used in this work is public available in [41].

**Acknowledgments:** This research has been partly supported by the Project: “Human-Robot Synergetic Logistics for High Value Crops” (project acronym: SYNERGIE), funded by the General Secretariat for Research and Technology (GSRT) under reference no. 2386.

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

## References

1. Sørensen, C.A.G.; Kateris, D.; Bochtis, D. ICT Innovations and Smart Farming. In Proceedings of the Communications in Computer and Information Science; Springer: Berlin/Heidelberg, Germany, 2019; Volume 953, pp. 1–19.
2. Moysiadis, V.; Tsolakis, N.; Katikaridis, D.; Sørensen, C.G.; Pearson, S.; Bochtis, D. Mobile Robotics in Agricultural Operations: A Narrative Review on Planning Aspects. *Appl. Sci.* **2020**, *10*, 3453. [[CrossRef](#)]
3. Nicolopoulou-Stamati, P.; Maipas, S.; Kotampasi, C.; Stamatis, P.; Hens, L. Chemical Pesticides and Human Health: The Urgent Need for a New Concept in Agriculture. *Front. Public Health* **2016**, *4*, 1. [[CrossRef](#)] [[PubMed](#)]
4. Kirkhorn, S.R.; Schenker, M.B. Current Health Effects of Agricultural Work: Respiratory Disease, Cancer, Reproductive Effects, Musculoskeletal Injuries, and Pesticide-Related Illnesses. *J. Agric. Saf. Health* **2002**, *8*, 199–214. [[CrossRef](#)] [[PubMed](#)]
5. Fathallah, F.A. Musculoskeletal disorders in labor-intensive agriculture. *Appl. Ergon.* **2010**, *41*, 738–743. [[CrossRef](#)]
6. Benos, L.; Tsaopoulos, D.; Bochtis, D. A Review on Ergonomics in Agriculture. Part II: Mechanized Operations. *Appl. Sci.* **2020**, *10*, 3484. [[CrossRef](#)]
7. Benos, L.; Tsaopoulos, D.; Bochtis, D. A review on ergonomics in agriculture. Part I: Manual operations. *Appl. Sci.* **2020**, *10*, 1905. [[CrossRef](#)]
8. Seay, J.F. Biomechanics of Load Carriage—Historical Perspectives and Recent Insights. *J. Strength Cond. Res.* **2015**, *29*, S129–S133. [[CrossRef](#)] [[PubMed](#)]
9. Fathallah, F.A.; Miller, B.J.; Miles, J.A. Low back disorders in agriculture and the role of stooped work: Scope, potential interventions, and research needs. *J. Agric. Saf. Health* **2008**, *14*, 221–245. [[CrossRef](#)] [[PubMed](#)]
10. Dahn, N.; Grass, H.M.; Fuchs, S. Situation Awareness for Autonomous Agents. In Proceedings of the RO-MAN 2018—27th IEEE International Symposium on Robot and Human Interactive Communication, Nanjing, China, 27–31 August 2018; pp. 666–671.
11. Salam, H.; Chetouani, M. A multi-level context-based modeling of engagement in human-robot interaction. In Proceedings of the 2015 11th IEEE International Conference and Workshops on Automatic Face and Gesture Recognition FG 2015, Ljubljana, Slovenia, 4–8 May 2015; Volume 2015-January.
12. Benos, L.; Bechar, A.; Bochtis, D. Safety and ergonomics in human-robot interactive agricultural operations. *Biosyst. Eng.* **2020**, *200*, 55–72. [[CrossRef](#)]
13. Yu, H.; Jia, W.; Li, Z.; Gong, F.; Yuan, D.; Zhang, H.; Sun, M. A multisource fusion framework driven by user-defined knowledge for egocentric activity recognition. *EURASIP J. Adv. Signal Process.* **2019**, *2019*, 14. [[CrossRef](#)]
14. Elijah, O.; Rahman, T.A.; Orikumhi, I.; Leow, C.Y.; Hindia, M.N. An Overview of Internet of Things (IoT) and Data Analytics in Agriculture: Benefits and Challenges. *IEEE Internet Things J.* **2018**, *5*, 3758–3773. [[CrossRef](#)]
15. Behera, A.; Hogg, D.C.; Cohn, A.G. Egocentric activity monitoring and recovery. In Proceedings of the Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), Shanghai, China, 6–9 June 2013; Volume 7726 LNCS, pp. 519–532.
16. Kitani, K.M.; Okabe, T.; Sato, Y.; Sugimoto, A. Fast unsupervised ego-action learning for first-person sports videos. In *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*; IEEE Computer Society: Piscataway, NJ, USA, 2011; pp. 3241–3248.
17. Zhan, K.; Faux, S.; Ramos, F. Multi-scale Conditional Random Fields for first-person activity recognition on elders and disabled patients. In *Proceedings of the Pervasive and Mobile Computing*; Elsevier B.V.: Amsterdam, The Netherlands, 2015; Volume 16, pp. 251–267.
18. Hoshen, Y.; Peleg, S. An Egocentric Look at Video Photographer Identity. In *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*; IEEE Computer Society: Piscataway, NJ, USA, 2016; Volume 2016-December, pp. 4284–4292.
19. Wang, L. Recognition of human activities using continuous autoencoders with wearable sensors. *Sensors* **2016**, *16*, 189. [[CrossRef](#)]
20. Nweke, H.F.; Teh, Y.W.; Al-garadi, M.A.; Alo, U.R. Deep learning algorithms for human activity recognition using mobile and wearable sensor networks: State of the art and research challenges. *Expert Syst. Appl.* **2018**, *105*, 233–261. [[CrossRef](#)]
21. Masum, A.K.M.; Bahadur, E.H.; Shan-A-Alahi, A.; Uz Zaman Chowdhury, M.A.; Uddin, M.R.; Al Noman, A. Human Activity Recognition Using Accelerometer, Gyroscope and Magnetometer Sensors: Deep Neural Network Approaches. In Proceedings of the 2019 10th International Conference on Computing, Communication and Networking Technologies, ICCCNT 2019, Kanpur, India, 6–8 July 2019.
22. Aguilera, A.A.; Brena, R.F.; Mayora, O.; Molino-Minero-Re, E.; Trejo, L.A. Multi-Sensor Fusion for Activity Recognition—A Survey. *Sensors* **2019**, *19*, 3808. [[CrossRef](#)]
23. Pham, M.; Yang, D.; Sheng, W. A Sensor Fusion Approach to Indoor Human Localization Based on Environmental and Wearable Sensors. *IEEE Trans. Autom. Sci. Eng.* **2019**, *16*, 339–350. [[CrossRef](#)]
24. Shoaib, M.; Bosch, S.; Incel, O.; Scholten, H.; Havinga, P. Fusion of Smartphone Motion Sensors for Physical Activity Recognition. *Sensors* **2014**, *14*, 10146–10176. [[CrossRef](#)]
25. Wu, W.; Dasgupta, S.; Ramirez, E.E.; Peterson, C.; Norman, G.J. Classification accuracies of physical activities using smartphone motion sensors. *J. Med. Internet Res.* **2012**, *14*. [[CrossRef](#)] [[PubMed](#)]
26. Agnusdei, G.P.; Elia, V.; Gnoni, M.G. A classification proposal of digital twin applications in the safety domain. *Comput. Ind. Eng.* **2021**, *154*, 107137. [[CrossRef](#)]

27. Lee, J.; Cameron, I.; Hassall, M. Improving process safety: What roles for Digitalization and Industry 4.0? *Process Saf. Environ. Prot.* **2019**, *132*, 325–339. [CrossRef]
28. Bulling, A.; Blanke, U.; Schiele, B. A tutorial on human activity recognition using body-worn inertial sensors. *ACM Comput. Surv.* **2014**, *46*. [CrossRef]
29. Liakos, K.; Busato, P.; Moshou, D.; Pearson, S.; Bochtis, D. Machine Learning in Agriculture: A Review. *Sensors* **2018**, *18*, 2674. [CrossRef]
30. Safi, K.; Mohammed, S.; Attal, F.; Khalil, M.; Amirat, Y. Recognition of different daily living activities using hidden Markov model regression. In *Proceedings of the Middle East Conference on Biomedical Engineering MECBME*; IEEE Computer Society: Piscataway, NJ, USA, 2016; Volume 2016–November, pp. 16–19.
31. Anguita, D.; Ghio, A.; Oneto, L.; Parra, X.; Reyes-Ortiz, J.L. Human activity recognition on smartphones using a multiclass hardware-friendly support vector machine. In *Proceedings of the Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*; Springer: Berlin/Heidelberg, Germany, 2012; Volume 7657 LNCS, pp. 216–223.
32. Shoaib, M.; Bosch, S.; Incel, O.; Scholten, H.; Havinga, P. Complex Human Activity Recognition Using Smartphone and Wrist-Worn Motion Sensors. *Sensors* **2016**, *16*, 426. [CrossRef] [PubMed]
33. Das, B.; Seelye, A.M.; Thomas, B.L.; Cook, D.J.; Holder, L.B.; Schmitter-Edgecombe, M. Using smart phones for context-aware prompting in smart environments. In *Proceedings of the 2012 IEEE Consumer Communications and Networking Conference, CCNC 2012, Las Vegas, NV, USA, 14–17 January 2012*; pp. 399–403.
34. Shoaib, M.; Bosch, S.; Incel, O.; Scholten, H.; Havinga, P. A Survey of Online Activity Recognition Using Mobile Phones. *Sensors* **2015**, *15*, 2059–2085. [CrossRef] [PubMed]
35. Milenkoski, M.; Trivodaliev, K.; Kalajdziski, S.; Jovanov, M.; Stojkoska, B.R. Real time human activity recognition on smartphones using LSTM networks. In *Proceedings of the 2018 41st International Convention on Information and Communication Technology, Electronics and Microelectronics, MIPRO 2018, Opatija, Croatia, 21–25 May 2018*; pp. 1126–1131.
36. Patil, P.A.; Jagyasi, B.G.; Raval, J.; Warke, N.; Vaidya, P.P. Design and development of wearable sensor textile for precision agriculture. In *Proceedings of the 2015 7th International Conference on Communication Systems and Networks, COMSNETS 2015, Bangalore, India, 29 August–5 September 2015*.
37. Sharma, S.; Raval, J.; Jagyasi, B. Mobile sensing for agriculture activities detection. In *Proceedings of the 3rd IEEE Global Humanitarian Technology Conference, GHTC 2013*; IEEE Computer Society: Piscataway, NJ, USA, 2013; pp. 337–342.
38. Sharma, S.; Raval, J.; Jagyasi, B. Neural network based agriculture activity detection using mobile accelerometer sensors. In *Proceedings of the 11th IEEE India Conference: Emerging Trends and Innovation in Technology, INDICON 2014*; Institute of Electrical and Electronics Engineers Inc.: Piscataway, NJ, USA, 2015.
39. Sharma, S.; Jagyasi, B.; Raval, J.; Patil, P. AgriAcT: Agricultural Activity Training using multimedia and wearable sensing. In *Proceedings of the 2015 IEEE International Conference on Pervasive Computing and Communication Workshops, PerCom Workshops 2015*; Institute of Electrical and Electronics Engineers Inc.: Piscataway, NJ, USA, 2015; pp. 439–444.
40. Bechar, A.; Vigneault, C. Agricultural robots for field operations: Concepts and components. *Biosyst. Eng.* **2016**, *149*, 94–111. [CrossRef]
41. Open Datasets—iBO. Available online: <https://ibo.certh.gr/open-datasets/> (accessed on 1 March 2021).
42. Reina, G.; Milella, A.; Galati, R. Terrain assessment for precision agriculture using vehicle dynamic modelling. *Biosyst. Eng.* **2017**, *162*, 124–139. [CrossRef]
43. Grimstad, L.; From, P.J. The Thorvald II Agricultural Robotic System. *Robotics* **2017**, *6*, 24. [CrossRef]
44. Lavender, S.A.; Li, Y.C.; Andersson, G.B.; Natarajan, R.N. The effects of lifting speed on the peak external forward bending, lateral bending, and twisting spine moments. *Ergonomics* **1999**, *42*, 111–125. [CrossRef] [PubMed]
45. Ghorri, G.M.U.; Luckwill, R.G. Responses of the lower limb to load carrying in walking man. *Eur. J. Appl. Physiol. Occup. Physiol.* **1985**, *54*, 145–150. [CrossRef] [PubMed]
46. Calibrate IMUs—IMU Documentation—Vicon Documentation. Available online: <https://docs.vicon.com/display/IMU/Calibrate+IMUs> (accessed on 7 January 2021).
47. Larsen, F.G.; Svenningsen, F.P.; Andersen, M.S.; de Zee, M.; Skals, S. Estimation of Spinal Loading During Manual Materials Handling Using Inertial Motion Capture. *Ann. Biomed. Eng.* **2020**, *48*, 805–821. [CrossRef]
48. Blue Trident IMU | Inertial Sensor by Vicon | Biomechanic Tracking. Available online: <https://www.vicon.com/hardware/blue-trident/> (accessed on 7 January 2021).
49. Burland, J.P.; Outerleys, J.B.; Lattermann, C.; Davis, I.S. Reliability of wearable sensors to assess impact metrics during sport-specific tasks. *J. Sports Sci.* **2020**. [CrossRef] [PubMed]
50. Garman, C.; Como, S.G.; Campbell, I.C.; Wishart, J.; O'Brien, K.; McLean, S. Micro-Mobility Vehicle Dynamics and Rider Kinematics during Electric Scooter Riding. In *Proceedings of the WCX SAE World Congress Experience 2020, Detroit, MI, USA, 21–23 April 2020*.
51. Tucker, H.W.; Tobin, E.R.; Moran, M.F. Tibial Accelerations During the Single-Leg Hop Test: Influence of Fixation. *J. Sport Rehabil.* **2020**, *1*, 1–4. [CrossRef] [PubMed]
52. Johnson, C.D.; Outerleys, J.; Tenforde, A.S.; Davis, I.S. A comparison of attachment methods of skin mounted inertial measurement units on tibial accelerations. *J. Biomech.* **2020**, *113*, 110118. [CrossRef] [PubMed]

53. Capture.U - IMeasureU. Available online: <https://imeasureu.com/capture-u/> (accessed on 19 January 2021).
54. iPad mini—Apple. Available online: <https://www.apple.com/ipad-mini/> (accessed on 19 January 2021).
55. Yu, H.; Pan, G.; Pan, M.; Li, C.; Jia, W.; Zhang, L.; Sun, M. A Hierarchical Deep Fusion Framework for Egocentric Activity Recognition using a Wearable Hybrid Sensor System. *Sensors* **2019**, *19*, 546. [[CrossRef](#)]
56. McGibbon, C.A.; Brandon, S.C.E.; Brookshaw, M.; Sexton, A. Effects of an over-ground exoskeleton on external knee moments during stance phase of gait in healthy adults. *Knee* **2017**, *24*, 977–993. [[CrossRef](#)]
57. Wang, Z.; Wu, L.; Sun, J.; He, L.; Wang, S.; Yang, L. Squat, stoop, or semi-squat: A comparative experiment on lifting technique. *J. Huazhong Univ. Sci. Technol. Med. Sci.* **2012**, *32*, 630–636. [[CrossRef](#)]
58. Vecchio, L. Del Choosing a Lifting Posture: Squat, Semi-Squat or Stoop. *MOJ Yoga Phys. Ther.* **2017**, *2*. [[CrossRef](#)]
59. Xiao, F.; Chen, J.; Xie, X.; Gui, L.; Sun, L.; Wang, R. SEARE: A System for Exercise Activity Recognition and Quality Evaluation Based on Green Sensing. *IEEE Trans. Emerg. Top. Comput.* **2020**, *8*, 752–761. [[CrossRef](#)]
60. Wang, L.; Gu, T.; Tao, X.; Lu, J. A hierarchical approach to real-time activity recognition in body sensor networks. *Pervasive Mob. Comput.* **2012**, *8*, 115–130. [[CrossRef](#)]
61. Bao, L.; Intille, S.S. Activity recognition from user-annotated acceleration data. In Proceedings of the Pervasive Computing, Second International Conference, PERVASIVE 2004, Vienna, Austria, 21–23 April 2004; pp. 1–17. [[CrossRef](#)]
62. sklearn.preprocessing.StandardScaler—scikit-learn 0.24. Available online: <https://scikit-learn.org/stable/modules/generated/sklearn.preprocessing.StandardScaler.html> (accessed on 20 January 2021).
63. Hochreiter, S.; Schmidhuber, J. Long Short-Term Memory. *Neural Comput.* **1997**, *9*, 1735–1780. [[CrossRef](#)] [[PubMed](#)]
64. Anagnostis, A.; Asimnari, G.; Papageorgiou, E.; Bochtis, D. A Convolutional Neural Networks Based Method for Anthracnose Infected Walnut Tree Leaves Identification. *Appl. Sci.* **2020**, *10*, 469. [[CrossRef](#)]
65. Demirkaya, A.; Chen, J.; Oymak, S. Exploring the Role of Loss Functions in Multiclass Classification. In Proceedings of the 2020 54th Annual Conference on Information Sciences and Systems, CISS 2020, Princeton, NJ, USA, 18–20 March 2020.
66. Xu, B.; Wang, N.; Chen, T.; Li, M. Empirical evaluation of rectified activations in convolutional network. *arXiv* **2015**, arXiv:1505.00853.
67. Goodfellow, I.; Bengio, Y.; Courville, A. 6.2.2.3 Softmax Units for Multinoulli Output Distributions. In *Deep Learning*; MIT Press: Cambridge, MA, USA, 2016; pp. 180–184, ISBN 978-0-26203561-3.
68. Shoaib, M.; Scholten, H.; Havinga, P.J.M. Towards physical activity recognition using smartphone sensors. In *UIC-ATC '13: Proceedings of the 2013 IEEE 10th International Conference on Ubiquitous Intelligence and Computing and 2013 IEEE 10th International Conference on Autonomic and Trusted Computing*; IEEE Computer Society: Washington, DC, USA, 2013; pp. 80–87.
69. Galin, R.; Meshcheryakov, R.; Kamesheva, S.; Samoshina, A. Cobots and the benefits of their implementation in intelligent manufacturing. In *Proceedings of the IOP Conference Series: Materials Science and Engineering*; Institute of Physics Publishing: Bristol, UK, 2020; Volume 862, p. 32075.
70. Ratsamee, P.; Mae, Y.; Kamiyama, K.; Horade, M.; Kojima, M.; Arai, T. Social interactive robot navigation based on human intention analysis from face orientation and human path prediction. *ROBOMECH J.* **2015**, *2*, 11. [[CrossRef](#)]
71. Charalampous, K.; Kostavelis, I.; Gasteratos, A. Recent trends in social aware robot navigation: A survey. *Rob. Auton. Syst.* **2017**, *93*, 85–104. [[CrossRef](#)]
72. Hameed, I.A.; Bochtis, D.D.; Sørensen, C.G.; Jensen, A.L.; Larsen, R. Optimized driving direction based on a three-dimensional field representation. *Comput. Electron. Agric.* **2013**, *91*, 145–153. [[CrossRef](#)]
73. Sørensen, C.G.; Bochtis, D.D. Conceptual model of fleet management in agriculture. *Biosyst. Eng.* **2010**, *105*, 41–50. [[CrossRef](#)]
74. Hameed, I.; Bochtis, D.; Sørensen, C. An Optimized Field Coverage Planning Approach for Navigation of Agricultural Robots in Fields Involving Obstacle Areas. *Int. J. Adv. Robot. Syst.* **2013**, *10*, 231. [[CrossRef](#)]
75. Kruse, T.; Pandey, A.K.; Alami, R.; Kirsch, A. Human-aware robot navigation: A survey. *Rob. Auton. Syst.* **2013**, *61*, 1726–1743. [[CrossRef](#)]
76. Marinoudi, V.; Sørensen, C.G.; Pearson, S.; Bochtis, D. Robotics and labour in agriculture. A context consideration. *Biosyst. Eng.* **2019**, *184*, 111–121. [[CrossRef](#)]
77. Sainath, T.N.; Vinyals, O.; Senior, A.; Sak, H. Convolutional, Long Short-Term Memory, fully connected Deep Neural Networks. In *Proceedings of the ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing*; Institute of Electrical and Electronics Engineers Inc.: Piscataway, NJ, USA, 2015; Volume 2015-August, pp. 4580–4584.
78. Shi, X.; Chen, Z.; Wang, H.; Yeung, D.-Y.; Wong, W.; Woo, W. Convolutional LSTM Network: A Machine Learning Approach for Precipitation Nowcasting. *Adv. Neural Inf. Process. Syst.* **2015**, *2015-January*, 802–810.