



Article

Human–Robot Interaction through Dynamic Movement Recognition for Agricultural Environments

Vasileios Moysiadis ^{1,2,3}, Lefteris Benos ¹, George Karras ², Dimitrios Kateris ¹, Andrea Peruzzi ⁴, Remigio Berruto ⁵, Elpiniki Papageorgiou ⁶ and Dionysis Bochtis ^{1,3,*}

¹ Institute for Bio-Economy and Agri-Technology (IBO), Centre of Research and Technology-Hellas (CERTH), 6th km Charilaou-Thermi Rd., 57001 Thessaloniki, Greece; v.moysiadis@farm-b.com (V.M.); e.benos@certh.gr (L.B.); d.kateris@certh.gr (D.K.)

² Department of Informatics and Telecommunications, University of Thessaly, 35100 Lamia, Greece; gkarras@uth.gr

³ farmB Digital Agriculture S.A., 17th November 79, 55534 Thessaloniki, Greece

⁴ Department of Agronomy and Agroecosystem Management, University of Pisa, Via S. Michele degli Scalzi 2, 56124 Pisa, Italy; andrea.peruzzi@unipi.it

⁵ Interuniversity Department of Regional and Urban Studies and Planning, University of Torino, Viale Mattioli 39, 10125 Torino, Italy; remigio.berruto@unito.it

⁶ Department of Energy Systems, University of Thessaly, Gaiopolis Campus, 41500 Larisa, Greece; elpinikipapageorgiou@uth.gr

* Correspondence: d.bochtis@certh.gr

Abstract: In open-field agricultural environments, the inherent unpredictable situations pose significant challenges for effective human–robot interaction. This study aims to enhance natural communication between humans and robots in such challenging conditions by converting the detection of a range of dynamic human movements into specific robot actions. Various machine learning models were evaluated to classify these movements, with Long Short-Term Memory (LSTM) demonstrating the highest performance. Furthermore, the Robot Operating System (ROS) software (Melodic Version) capabilities were employed to interpret the movements into certain actions to be performed by the unmanned ground vehicle (UGV). The novel interaction framework exploiting vision-based human activity recognition was successfully tested through three scenarios taking place in an orchard, including (a) a UGV following the authorized participant; (b) GPS-based navigation to a specified site of the orchard; and (c) a combined harvesting scenario with the UGV following participants and aid by transporting crates from the harvest site to designated sites. The main challenge was the precise detection of the dynamic hand gesture “come” alongside navigating through intricate environments with complexities in background surroundings and obstacle avoidance. Overall, this study lays a foundation for future advancements in human–robot collaboration in agriculture, offering insights into how integrating dynamic human movements can enhance natural communication, trust, and safety.

Keywords: human–robot collaboration; natural communication framework; vision-based human activity recognition; situation awareness



Citation: Moysiadis, V.; Benos, L.; Karras, G.; Kateris, D.; Peruzzi, A.; Berruto, R.; Papageorgiou, E.; Bochtis, D. Human–Robot Interaction through Dynamic Movement Recognition for Agricultural Environments.

AgriEngineering **2024**, *6*, 2494–2512.
<https://doi.org/10.3390/agriengineering6030146>

Academic Editors: Marcello Biocca and Massimo Cecchini

Received: 28 June 2024

Revised: 24 July 2024

Accepted: 29 July 2024

Published: 1 August 2024



Copyright: © 2024 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

Robots are a key component of Agriculture 4.0, where advanced technologies are revolutionizing traditional farming practices. Robotic systems can significantly improve agricultural productivity by enhancing the efficiency of several operations, allowing farmers to make more informed decisions and implement targeted interventions [1,2]. In addition, agri-robots offer solutions to labor shortages for seasonal work and can perform tasks in hazardous environments, reducing human risk [3,4]. Typically, robots perform programmed actions based on limited and task-specific instructions. However, unlike the

controlled environments of industrial settings with well-defined objects, open-field agriculture involves unpredictable and diverse conditions [5,6]. As a consequence, agri-robots must navigate complex environments with varying physical conditions as well as handle live crops delicately and accurately. Factors such as lighting, terrain, and atmospheric conditions are inconsistent, while there is a substantial variability in crop characteristics like shape, color, and position, which are not easily predetermined. These complexities make the substitution of human labor with autonomous robots in agriculture a considerable challenge.

To address the challenges presented by complex agricultural environments, human–robot collaboration has been proposed to achieve shared goals through information exchange and optimal task coordination. This multidisciplinary research field, usually called human–robot interaction (HRI), is continuously evolving because of the rise of information and communications technologies (ICTs), including artificial intelligence (AI) and computer vision to improve robot perception, decision making, and adaptability [7–11]. HRI merges also elements of ergonomics [12] as well as social [13] and skill development sciences [14] to ensure that human–robot collaboration is not only efficient but also safe and socially acceptable. HRI leverages human cognitive skills, such as judgment, dexterity and decision making, alongside the robots' strengths in repeatable accuracy and physical power. These semi-autonomous systems have shown superior performance compared to fully autonomous robots. In short, human–robot synergy offers numerous benefits, including increased productivity, flexibility in system reconfiguration, reduced workspace requirements, rapid capital depreciation, and the creation of medium-skilled jobs [15,16].

Recent review studies on HRI in agriculture, such as [15,17–20], highlight that efficient human–robot collaboration necessitates the creation of natural communication frameworks to facilitate knowledge sharing between humans and robots. This involves exploring communication methods that mirror those used in everyday life, aiming for a more intuitive interaction between humans and robots. An ideal approach would be vocal communication in terms of short commands or dialogues [21]. Nevertheless, this approach is constrained by the ambient noise levels prevalent in agricultural environments and variations in pronunciation among individuals [22,23]. Typing commands would also be precise enough, but impractical in dynamic agricultural settings, where rapid and fluid HRI is needed. Users generally prefer more natural forms of communication over typing commands [24]. On the other hand, leveraging image-based recognition for interpreting human movements provides a seamless and adaptable means of interaction that can enhance the efficiency and effectiveness of human–robot collaboration in agriculture. However, challenges still remain concerning mainly varying environmental conditions and potential visual obstructions, which can complicate the effectiveness of visual recognition [25].

Body language, such as body postures and hand gestures, seems to be a promising alternative solution. Despite limited research, promising results have been accomplished in spite of challenges such as constraints on human movements, the production of extensive and noisy sensor datasets, and obstacles encountered by vision sensors in constantly changing agricultural environments and fluctuating lighting conditions. Indicatively, wearable sensors have been employed in [26] for predicting predefined human activities for an HRI scenario involving two different unmanned ground vehicles (UGVs) aiding workers during harvesting tasks. Remarkably, UGVs have demonstrated their suitability as co-workers for agricultural applications by meeting essential criteria, such as compact size, low weight, advanced intelligence, and autonomous capabilities [27–29].

Hand gesture recognition, facilitated either by sensors embedded within gloves [30] or vision sensors [25], has also attracted attention in agriculture. In the former study [30], gestures using a hand glove were utilized to control wirelessly (via Bluetooth) a robotic arm for performing weeding operations. In [25], coordination between a UGV and workers was accomplished, as the UGV followed the workers during harvesting and facilitated crate transportation from the harvesting site to a designated location. This involved the development of a hand gesture recognition framework by integrating depth videos and

skeleton-based recognition. Specifically, Moysiadis et al. [25] investigated converting five usual static hand gestures into corresponding explicit UGV actions. However, static gestures may not capture the full range of human movement and intention, limiting their versatility in dynamic situations and leading to potential misinterpretation or ambiguity in communication. In contrast, dynamic movements have the potential to offer greater flexibility and richer contextual information by incorporating a broader range of body movements and variations, enabling more nuanced and natural interactions between humans and robots [31].

As a means of enhancing natural communication during HRI, the present AI-assisted framework incorporated vision-based dynamic human movement recognition. Furthermore, safety measures were taken into account to ensure that HRI does not compromise the well-being of humans involved in the interaction. Combinations of the developed capabilities of the proposed HRI system were field tested in three experimental scenarios to assure its adaptability to relevant situations taking place in open-field agricultural settings. The main challenges encountered during these tests were thoroughly identified and analyzed, providing valuable insights for future improvements in the system's design and functionality. To the authors' knowledge, this is the first verification of such an interactive system in a real orchard environment.

The remainder of the present paper is structured as follows: Section 2 presents the materials and methods, providing information about the overall system architecture, vision-based dynamic movement recognition framework (data acquisition and processing, tested machine learning (ML) algorithms, allocation of human movements to corresponding UGV actions), UGV autonomous navigation as well as the examined field scenarios including the utilized hardware and software. Section 3 delves into the results regarding the performance of the examined ML algorithms and the implementation of the developed AI-assisted HRI framework in an orchard environment. Section 4 provides concluding remarks, accompanied by a discussion within a broader context, along with suggestions for future research directions.

2. Materials and Methods

2.1. Overall System Architecture

In Figure 1, a schematic of the data flow for the proposed framework is presented. In summary:

- The system exploits video frames as data input captured by a depth camera mounted on a UGV. These frames are transferred via a Robot Operating System (ROS) topic to an ROS node responsible for extracting the red–green–blue (RGB) and depth values from each frame. This separation of color information from depth facilitates analysis and preparation for the pose detection algorithm.
- As a next step, the OpenPose algorithm node is used, which takes both the RGB and depth channels as inputs in order to extract essential body joints from the video frames.
- The extracted body joints from the OpenPose algorithm are then fed into the ML algorithm for movement detection after passing the required data preprocessing phase.
- To communicate with the rest of the system, various ROS nodes were developed requiring the output of the ML algorithm.
- The ROS nodes are classified into two interconnected categories related to decision and action:
 - (1) The first category is responsible for the decision-making process based on the output of the ML algorithm. These nodes interpret the movement recognition results and make decisions depending on the scenario. These decisions could involve controlling the robot or triggering specific actions based on the recognized movements. Throughout the entire data flow, communication between different nodes occurs through ROS topics and services. In general, ROS topics facilitate the exchange in data between nodes, while ROS services enable the

request–response patterns, allowing nodes to request specific information or actions from other nodes [25].

- (2) The second category of nodes addresses UGV actions. These nodes are triggered by the decision-making nodes and integrate data from the Navigation Stack [32], sensor values, and key performance indicators (KPIs) related to the progress of each operation. Furthermore, the second category of nodes provides feedback to the decision-making nodes, ensuring that the UGV's actions are continuously monitored and adjusted based on the feedback.

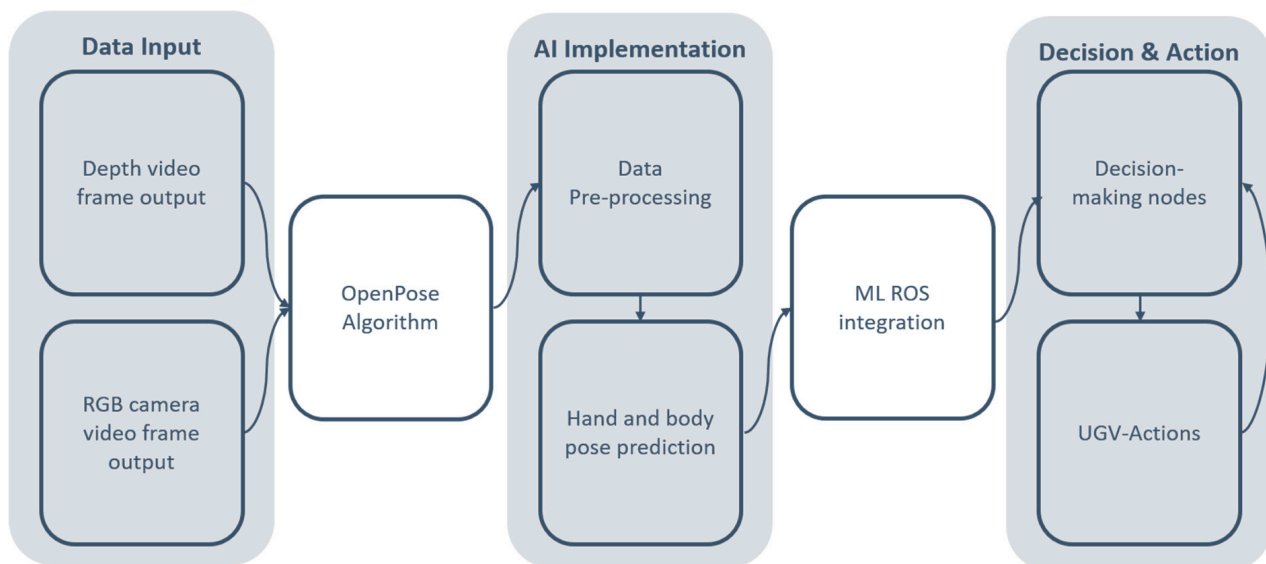


Figure 1. Data flow diagram of the proposed framework.

2.2. Movement Recognition

2.2.1. Data Input

The data acquisition process was performed by a depth camera mounted on an agricultural robot capable of autonomous operation. Ten participants with varying anthropometric characteristics participated in the experimental sessions. To further increase variability, the distance between the camera and each participant was intentionally varied, while the participants performed multiple iterations and variations of each movement. All camera shots were captured on both sunny and cloudy days at different locations within an orchard situated in the Thessaly region of central Greece, enriching the dataset's diversity. The camera was set to record at 60 frames per second, as in [25], assuring a high level of precision in capturing the movements of both hands and the participants' body posture.

The examined dynamic hand gestures, which are depicted in Figure 2a–g, are commonly classified into the following categories:

- “Left hand waving”: a repetitive leftward and rightward motion of the left hand;
- “Right hand waving”: a repetitive leftward and rightward motion of the right hand;
- “Both hands waving”: a repetitive leftward and rightward motion of both hands simultaneously;
- “Left hand come”: a beckoning motion with the left hand, moving toward the body;
- “Left hand fist”: the motion of clenching the left hand into a fist;
- “Right hand fist”: the motion of clenching the right hand into a fist;
- “Left hand raise”: an upward motion of the left hand, lifting it above shoulder level.



Figure 2. Depiction of the examined dynamic hand gestures: (a) “left hand waving”, (b) “right hand waving”, (c) “both hands waving”, (d) “left hand come”, (e) “left hand fist” (f) “right hand fist”, and (g) “left hand raise”.

In addition to the above hand movements, two predefined human activities were added in the training procedure for enhancing situation awareness:

- “Standing”: the participant remains stationary in an upright position;

- “Walking”: the dynamic activity of the participant moving randomly in a walking motion.

2.2.2. Prediction of Body Joint Positions

As a next step, the OpenPose algorithm was implemented, which is an ML-based approach for estimating human poses by accurately predicting the positions of body joints [33]. In particular, OpenPose belongs to Convolutional Neural Networks (CNNs), which is a form of deep learning model noted for its ability to extract meaningful characteristics from images. Multiple layers, including convolutional, pooling and fully connected layers, are used to train the model toward learning hierarchical representations of the input images. A vast collection of labeled human pose instances is fed into the network throughout the training process. The network is trained to optimize its parameters by reducing the discrepancy between anticipated and true joint locations. OpenPose refines initial estimates of joint positions using a cascaded technique. The network’s first stage produces a rough estimation of the joint positions. This estimation is then refined in stages with each stage focusing on improving the accuracy of the prior estimation. This cascaded architecture enables the network to refine the pose estimation and collect finer features over time.

An algorithm was developed using Python and the ROS framework that implements OpenPose for movement recognition. In Figure 3, a graph is shown illustrating the data flow from the RGB-D image captured by the depth camera to the dedicated node responsible for consolidating training and testing datasets. In short, the algorithm used an ROS node (OpenPose node) to subscribe to the image topic (cloud registered) in order to retrieve the point cloud frame. After the frame was received, the OpenPose algorithm was applied. The produced results were published via an ROS topic (detected people landmarks) using a custom ROS message. Another ROS node (train registered) was developed to aggregate the above-mentioned results in order to create the train–testing datasets.

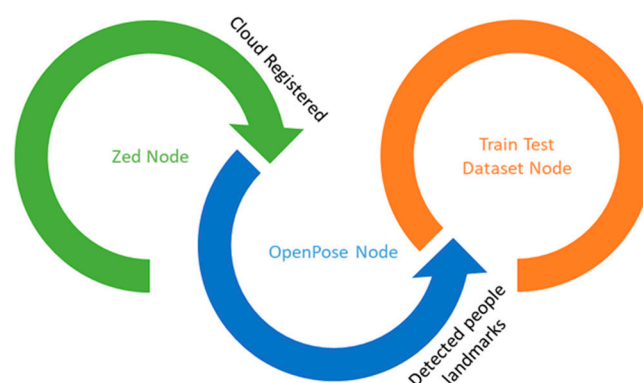


Figure 3. Chart depicting the data flow between the captured RGB-D images to the developed node, which is responsible for aggregating the training/testing data.

Next, for the purpose of aggregating the training and testing data, a dedicated ROS node was created. This node subscribes to the generated results topic. There are two main constants that have to be predefined. The first constant is related to the data acquisition time ($t_{data\ acquisition}$) and was set equal to 120 sec so as to ensure class balance. Hence, all the movements should be executed using the same time duration. The second constant, namely movement time duration ($t_{movement}$) was set equal to 1 s and is related to the time that a person should continuously perform each movement so as to properly be distinguished from random movement.

When the callback function is activated, the developed algorithm using specific body reference landmarks computes the relative angles of both hands in relation to the torso (the axial part of the human body) ($\hat{\xi}_{left\ torso}$ and $\hat{\delta}_{right\ torso}$), the shoulders ($\hat{\kappa}_{left\ shoulder}$ and $\hat{\lambda}_{right\ shoulder}$), the elbows ($\hat{\theta}_{left\ elbow}$ and $\hat{\varphi}_{right\ elbow}$), and the chest ($\hat{\mu}_{left\ chest}$ and $\hat{\nu}_{right\ chest}$). The aforementioned anthropometric parameters are presented in Figure 4.



Figure 4. Depiction of the examined anthropometric parameters, namely the relative angles of both hands in relation to (a) the shoulders ($\hat{\kappa}_{left\ shoulder}$ and $\hat{\lambda}_{right\ shoulder}$), the elbows ($\hat{\theta}_{left\ elbow}$ and $\hat{\phi}_{right\ elbow}$), the chest ($\hat{\mu}_{left\ chest}$ and $\hat{\nu}_{right\ chest}$) and (b) the torso ($\hat{\xi}_{left\ torso}$ and $\hat{\delta}_{right\ torso}$).

2.2.3. Class Balance

The present dataset used for the training process contains approximate 9000 iterations per class. The dataset’s rows correspond to the number of frames within $t_{movement}$, whereas the columns represent the values of anthropometric angles of the frame presented in Figure 4. During the training of an ML algorithm, it is critical to use a dataset with nearly equal numbers of samples for each class, since class imbalance may lead to misleading performance metrics and other problems. In the current experimental sessions, the participants intentionally carried out the prescribed movements within the same time window ($t_{data\ acquisition} = 120\ s$). Consequently, as shown in Figure 5, the classes are well balanced.

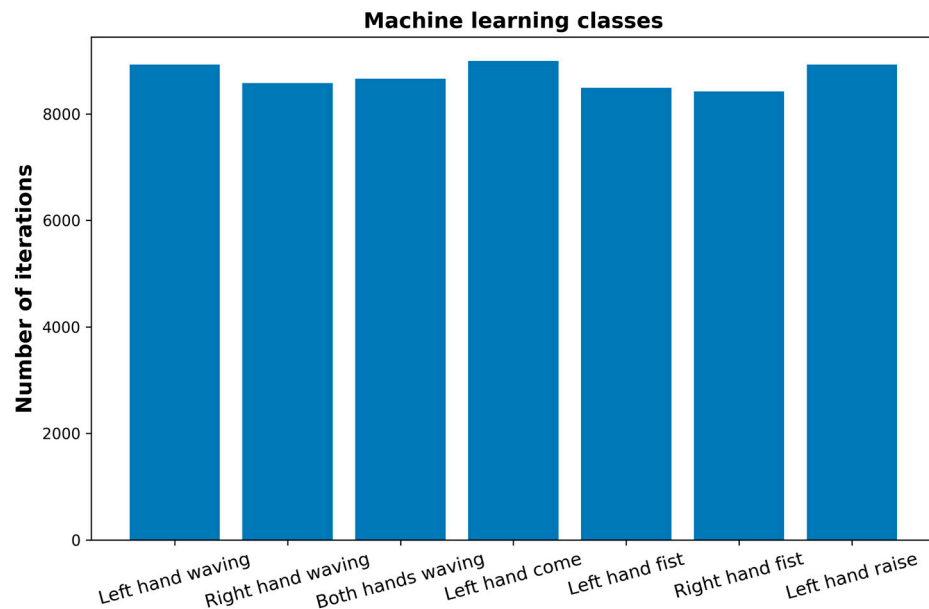


Figure 5. Bar chart illustrating the number of instances of each class with each bar corresponding to a specific dynamic hand gesture.

2.2.4. Tested Machine Learning Algorithms

ML classifiers use input training data to predict the likelihood that data will fall into one of these classifications. Numerous ML algorithms are well reported in the literature [34–37]. The following algorithms were tested in this analysis, which are usually used in such kinds of problems:

- Logistic Regression (LR): This method estimates discrete values by analyzing independent variables, allowing the prediction of event likelihood by fitting data to a logit function.
- Linear Discriminant Analysis (LDA): LDA projects features from a higher-dimensional space to a lower-dimensional one, reducing dimensions without losing vital information.
- K-Nearest Neighbor (KNN): KNN is a pattern recognition algorithm that identifies the closest neighbors for future samples based on training datasets, making predictions according to similar instance characteristics.
- Classification and Regression Trees (CART): This tree-based model operates on “if-else” conditions, where decisions are made based on input feature values.
- Naive Bayes: Naive Bayes is a probabilistic classifier that assumes feature independence within a class, treating them as unrelated when making predictions.
- Support Vector Machine (SVM): SVM maps raw data points onto an n-dimensional space, facilitating effective data classification based on their spatial distribution in the transformed space.
- LSTM (Long Short-Term Memory): LSTM is a neural network architecture designed for processing sequential input. It addresses typical recurrent neural network (RNN) limitations by incorporating memory cells with input, forget, and output gates to selectively store, forget, and reveal information. This enables LSTM networks to capture long-term dependencies and retain contextual information across extended sequences. LSTM optimizes memory cells and gate settings through training via backpropagation through time (BPTT), making it highly effective for tasks requiring knowledge of sequential patterns and relationships.

2.2.5. Optimizing LSTM-Based Neural Network Model for Multiclass Classification: Architecture, Training, and Performance Enhancement

The model fine tuning in the training phase of the network involves several steps to maximize its performance. In particular, the architecture of the model has three LSTM layers that are composed of 64, 128, and 64 units, respectively, each with the hyperbolic tangent (tanh) activation function. The tanh activation function was initially selected over the Rectified Linear Unit (ReLU) function, because it more effectively addresses the vanishing gradient problem [38]. By using dropout layers of 0.2 after the first two LSTM layers, the overfitting issue was solved as the input units were set equal to zero with a given fraction during the training. Then, the final dense layers were added, including 64 and 32 units, with the ReLU along with the softmax activation function to the output layer being used for multiclass classification. To compile the model, the Adam optimizer was chosen due to its adaptive learning rate characteristics, while the learning rate was set equal to 0.001. The loss function for this was “*categorical_crossentropy*”, while the performance metric was “*categorical_accuracy*”, which calculates how many samples in the set are correctly classified. The training was assisted by two callbacks, namely “*EarlyStopping*” and “*ReduceLROnPlateau*”. The former is used to observe the validation loss and terminates the training process if 20 consecutive epochs elapse with no improvement. It also returns the prevailing weights to prevent overfitting. The “*ReduceLROnPlateau*” decreases the learning rate by a factor of 0.5 if there are 10 epochs without improvement in validation loss. The model was trained for a maximum of 500 epochs with a cross-validation split of 25%, enabling the model to generalize on unknown data.

2.3. Allocation of Human Movement Detection

In this subsection, the allocation of specific human movements to corresponding actions of the UGV, to facilitate robust HRI in various scenarios, is presented. In brief, a UGV can identify a participant requiring assistance and lock onto them using the “left hand waving” gesture. Once locked, this person becomes the sole authorized individual to interact with the UGV using particular commands, minimizing potential safety concerns [12,39].

The locked participant can make use of six more commands that are associated with distinct hand gestures and the corresponding UGV's actions. Specifically, the "right hand waving" gesture instructs the UGV to autonomously navigate to a predefined location, while the "left hand come" gesture indicates the UGV to follow the participant. The HRI system also includes gestures for pausing specific operations. The "left hand fist" gesture temporarily stops the UGV from following the participant, whereas the "right hand fist" gesture commands the UGV to return to the site where autonomous navigation was initially activated. At any time, the participant can be unlocked by using the "left hand raise" gesture, allowing another individual to take control if necessary. Finally, the "both hands waving" movement serves as an emergency stop signal, prompting the UGV to halt all actions and distance itself from nearby humans and static obstacles to ensure safety.

Furthermore, the UGV can detect the participant's body posture, distinguishing between an upright stance ("standing") and walking activity ("walking"). This capability enhances situational awareness, allowing the UGV to adapt to dynamic changes in the environment. In turn, when humans recognize the robot's environmental awareness, they can build trust and confidently delegate tasks that the robot can effectively complete [40]. Table 1 provides a summary of all the examined gestures, each associated with a specific class (used for ML classifiers) and their corresponding UGV actions.

Table 1. Summary of the examined hand and body movements along with the corresponding classes and actions to be executed by the UGV.

Recognized Human Movement	Corresponding Class	UGV Action
"Right hand waving"	0	Initiate autonomous navigation to a predefined location
"Left hand waving"	1	Lock identified person
"Both hand waving"	2	Stop all active scenarios
"Left hand come"	3	Follow the participant
"Left hand fist"	4	Pause following mode
"Right hand fist"	5	Return to the location that the autonomous navigation was initiated
"Left hand raise"	6	Unlock identified participant
"Standing"	7	Detect upright posture
"Walking"	8	Detect walking activity

The dynamic hand gestures examined in this study were used for enabling versatile and responsive interaction between participants and UGVs in an orchard setting. These gestures were designed to offer users a range of options without causing confusion, as not all gestures are intended for simultaneous use. This communication framework addresses a variety of operational scenarios, allowing the UGV to effectively manage tasks like autonomous navigation, halting actions, following participants, and handling emergency situations. Hence, the UGV can adapt to the diverse needs of the orchard environment, enhancing operational efficiency and worker safety through precise and intuitive interaction.

2.4. Autonomous Navigation

The implemented UGV (Thorvald, SAGA Robotics SA, Oslo, Norway), besides being equipped with the RGB-D camera (Stereolabs Inc., San Francisco, CA, USA) for human movement recognition, featured several sensors to support the agricultural application. Specifically, a laser scanner (Velodyne Lidar Inc., San Jose, CA, USA) was utilized to scan the environment and create a two-dimensional map. An RTK GPS (S850 GNSS Receiver, Stonex Inc., Concord, NH, USA) and inertial measurement units (RSX-UM7 IMU, RedShift Labs, Studfield, VIC, Australia) provided velocity, positioning (latitude and longitude coordinates), and time information to enhance robot localization, navigation, and obstacle avoidance. To meet the computational demands, all vision-related operations were

managed by a Jetson TX2 Module with CUDA support. Figure 6 shows the UGV platform used in this study, which was equipped with the essential sensors.



Figure 6. The unmanned ground vehicle utilized in the present study along with the embedded sensors.

The commands related to the UGV's navigation were executed via a modified version of the ROS Navigation Stack [41–43]. As a global planner, the modified version of the NavFn [44] was implemented to enable global localization and GPS-based navigation. Furthermore, the following local planners for obstacle avoidance were tested in the field under similar environmental conditions [45,46]:

- Trajectory Rollout Planner (TRP): Ideal for outdoor settings, the TRP utilizes kinematic constraints to provide smooth trajectories. This planner effectively manages uneven terrain and various obstacles by evaluating different paths and selecting the one that avoids obstacles while maintaining smooth motion.
- TEB Planner: An extension of the EBand method, the TEB Planner considers a short distance ahead of the global plan and generates a local plan comprising a series of intermediate UGV positions. It is designed to work with dynamic obstacles and kinematic limitations common in outdoor environments. By optimizing both local and global paths, it balances obstacle avoidance with goal achievement. This planner's ability to handle dynamic obstacles is particularly useful in outdoor settings where objects such as workers, livestock, animals, and agricultural equipment can move unpredictably.
- DWA Planner: While the DWA Planner is effective for dynamic obstacle avoidance, it may not handle complex outdoor terrains as efficiently as the TRP or TEB Planner. It focuses primarily on avoiding moving obstacles and might not perform as well in intricate outdoor scenarios.

After extensive testing of the above planners in field conditions, it was determined that the TEB Planner was the most suitable for the present HRI system. Remarkably, the efficiency of planners such as the TRP and DWA is influenced by factors like sensor capabilities, map quality, and unpredictable environmental conditions. Considering these factors, the TEB Planner was chosen for its superior performance in handling dynamic obstacles and kinematic constraints common in dynamic agricultural environments. The TEB Planner was fine tuned with numerous parameters, including adjustments to the local cost map based on the physical limitations of the employed UGV, to facilitate optimal autonomous navigation.

2.5. Brief Description of the Implemented Scenarios

In this section, three distinct scenarios are shortly described, which were designed to test and validate the effectiveness of the present HRI framework. An attempt was made to emulate various real-world outdoor agricultural tasks and conditions, providing useful insights into the system's capabilities and limitations. Each scenario focuses on different aspects of the UGV's functionality. By implementing these diverse scenarios, we aim to demonstrate the robustness and versatility of the developed HRI system in enhancing agricultural productivity, flexibility and safety. Toward that direction, a thorough step-by-step approach was adopted to rigorously test the capabilities of the developed system.

2.5.1. UGV Following Participant at a Safe Distance and Speed

Field tests were conducted involving either one or two individuals interacting with the UGV as a means of demonstrating the capability of the UGV to distinguish between the two participants and identify which person required assistance and respond accordingly. Actions included also following the person who signaled for assistance, halting its movement, returning to a predetermined location, and unlocking the participant.

2.5.2. GPS-Based UGV Navigation to a Predefined Site

In this scenario, the UGV autonomously navigated to a designated location. Utilizing GPS coordinates, the UGV calculated the optimal route and traversed outdoor terrain, adjusting its path as needed to avoid obstacles. Upon reaching the location, it slowly approached the predefined site, showcasing its ability to navigate effectively in real-world agricultural environments.

2.5.3. Integrated Harvesting Scenario Demonstrating All the Developed HRI Capabilities

In this scenario, the UGV was programmed to follow participants at a safe distance and speed during harvesting activities. Moreover, it was tasked with approaching authorized persons to assist in the transportation of crates from the harvesting site to a designated location. The UGV's functionality also included the ability to detect and respond to emergency situations. For instance, it could promptly respond to commands issued by participants for immediate halting all its operation. This all-inclusive scenario allowed for the evaluation of various aspects of the UGV performance, including navigation, obstacle avoidance, interaction with human workers, and emergency response capabilities.

3. Results

3.1. Machine Learning Algorithms Performance Comparison for Classification of Human Movements

As analyzed in Section 2.2.4, seven ML algorithms, namely LR, LDA, KNN, CART, NB, SVM, and LSTM, were examined to ascertain their ability to predict the dynamic hand and body movements listed in Table 1. Each algorithm underwent rigorous testing to evaluate its effectiveness in classifying these movements into the corresponding classes. For this purpose, the classification report depicted in Table 2 is used to provide insights into the precision, recall, and F1-score for each class, offering valuable information into each ML models' strengths and weaknesses. As stressed in the methodology section, the investigated nine classes were deliberately balanced during preprocessing.

In summary, NB and CART showed the lower precision, recall and F1-score values, suggesting challenges in accurately classifying the examined movements with these algorithms. In contrast, LR, LDA, KNN, SVM, and LSTM exhibited relatively high scores, indicating robust performance in correctly identifying movements and minimizing false positives and negatives. Among the latter group of ML algorithms, LSTM emerged as the best classifier, demonstrating the highest overall average score. Hence, LSTM was selected to be the ML classifier of the proposed HRI framework with technical information relating to architecture, training, and performance enhancement provided in Section 2.2.5. Interestingly, our analysis revealed that the "left hand come" dynamic hand gesture presents

the utmost challenges for accurate detection. In fact, the “come” gesture involves the hand positioned in front of the human body, making it difficult to find specific reference landmarks for recognition. The dynamic nature of this hand movement, coupled with variations in arm positions and anthropometric features, adds complexity to the recognition process. These challenges demonstrated the need for robust ML algorithms and sensor capabilities to ensure accurate recognition and timely response [47].

Table 2. Classification report for the performance of the seven investigated machine learning algorithms.

Algorithm	Precision	Recall	F1-Score
LR	0.866	0.863	0.862
LDA	0.9135	0.883	0.879
KNN	0.871	0.842	0.838
CART	0.765	0.698	0.686
NB	0.108	0.328	0.162
SVM	0.871	0.842	0.838
LSTM	0.912	0.95	0.90

3.2. System Prototype Demonstration in Different Field Scenarios

In Section 2.5, we briefly described the methodology for testing the developed HRI framework through three carefully designed scenarios. The results of these field tests are presented in this section. Each scenario was tailored to assess the specific functionalities of the UGV, including navigation, obstacle avoidance, human movement recognition, situation awareness, and emergency response capabilities. Through rigorous testing and analysis, valuable understandings were gained into the system’s performance and its potential for building trust, thus ensuring reliable and efficient task delegation.

It should be stressed that the robotic vehicle moved autonomously in the field, by always maintaining a predefined safe distance from individuals. This distance was set to be at least 0.5 m, with a maximum limit of 15 m (due to sensor restrictions), to ensure social navigation. Unlike purely autonomous navigation, social navigation inspires a sense of safety in workers, enhancing their comfort and confidence in the robotic system [12]. To guarantee that the above limits were maintained, both a minimum distance and a maximum speed of 0.2 m/s from identified individuals and potential obstacles were programmed, constituting necessary prerequisites. Similarly to [25], the open-source package “ROS Navigation Stack” was utilized. In all scenarios, the UGV could detect the “standing” and “walking” activity, allowing, as mentioned above, the robotic system to adjust to evolving conditions in the environment.

Subsequently, indicative images taken during the experimental sessions are provided for each scenario, while more images depicting also the examined dynamic human movements are presented in the final integrated harvesting scenario.

3.2.1. UGV Following Participant at a Safe Distance and Speed

A methodical approach was adopted, starting with the evaluation of the proposed framework in a scenario involving a single individual working alongside the UGV. As the next step, also a second person was added in the same working area with the UGV simulating practical scenarios encountered in agricultural operations. In the two-person pilot setup, the UGV successfully showcased its ability to distinguish between the two participants and determine which one needed assistance by detecting the corresponding movement assigned to locking. Depending on the specific human movement detected, the UGV executed the corresponding actions, including following the person signaling for assistance, pausing the following mode for a while, stopping all UGV active operations, returning to a predefined area, or unlocking the participant.

The human movements designated for this scenario included the subsequent human movements: (a) “left hand waving” to lock the person performing the specific dynamic gesture; (b) “left hand come” to initiate the following mode; (c) “left hand fist” to pause the

following mode; (d) “left hand raise” for the locked individual to request to be unlocked by the UGV; (e) “both hand waving” to stop all active operations of the UGV; and (f) “right hand fist” to return to a predefined site located within the orchard.

In summary, the implementation of this field scenario proved to be successful. By accurately recognizing and responding to all the above dynamic gestures, the UGV kept also the aforementioned prerequisites regarding the safe limits of speed and distance. Additionally, the UGV immediately stopped once the corresponding hand gesture was detected. Figure 7a depicts the UGV following the participant, while Figure 7b illustrates the participant requesting the UGV to pause the following mode by showing his left fist.



Figure 7. Images from the experimental demonstration of the “following” scenario in an orchard showing (a) the UGV following the locked participant while maintaining a safe speed and distance and (b) the participant asking the UGV to pause the following mode.

A notable challenge emerged during the testing phase, necessitating further refinement of the algorithm. This challenge concerned ensuring reliable gesture recognition across different environmental conditions, such as varying lighting and background complexity. Therefore, we purposely conduct this scenario on several sites of the orchard during both sunny and cloudy days. In this way, this challenge was overcome resulting in the successful execution of this field scenario, demonstrating the potential of the developed HRI framework to support the “following” mode in open-field agricultural settings.

3.2.2. GPS-Based UGV Navigation to a Predefined Site

In this scenario, the UGV autonomously navigated to a designated location. Leveraging GPS coordinates, the UGV computed the most efficient route and traversed outdoor terrain, making necessary adjustments to avoid obstacles along the way. Upon reaching the location, the UGV safely approached the predefined site, demonstrating its capability to navigate effectively in real-world agricultural environments. It should be mentioned that if the target is determined to be unreachable due to obstacles or excessive distance, the UGV halts its operation and waits for further instructions from the designated participant. This protocol ensures that the UGV maintains safe and reliable operation in all scenarios.

A methodical approach was undertaken again, beginning with the evaluation of the proposed framework in a setting involving only the UGV. Subsequently, one and two participants were gradually added in the same region with the UGV to simulate practical scenarios. In the two-person setup, the UGV successfully exhibited, again, its ability to distinguish between the two participants and determine which one required assistance by detecting the corresponding movement assigned to locking. The human movements designated for this scenario included (a) “left hand waving” for locking the authorized person; (b) “left hand raise” for unlocking; (c) “right hand waving” for initiating the mode of autonomous navigation to the predefined location; and (d) “both hand waving” for stopping all UGV operations.

The implementation of this scenario was successful, demonstrating the UGV's ability to navigate to the predefined site autonomously, as can be seen in Figure 8. A noteworthy challenge that emerged during testing related to ensuring who would be the authorized person was again overcome, while all obstacles were identified by the UGV enabling smooth navigation within the orchard.



Figure 8. Image from the UGV autonomously navigating within the orchard toward a predefined site indicated by the map location sign marker.

3.2.3. Integrated Harvesting Scenario Demonstrating All the Developed HRI Capabilities

This scenario was the most challenging one, as it was designed to combine all the capabilities of the developed HRI framework in an open-field agricultural environment. One of the primary challenges encountered was the integration of multiple tasks, including autonomous navigation, distance maintenance from participants, obstacle avoidance, discerning which participant should be locked by the UGV, and emergency response, into a cohesive and fluid workflow. In spite of these challenges, the system demonstrated remarkable adaptability and responsiveness. Its success was also assured by the step-by-step methodological approach, where combinations of HRI capabilities had already been assessed in different situation-specific scenarios.

Figure 9 depicts some images taken during the experimental demonstration of the predesigned harvesting scenario, illustrating how the UGV smoothly transitioned between different modes of operation. Specifically, in Figure 9a, the locked participant asks the UGV to immediately stop all active scenarios via the “both hand waving” movement, while in Figure 9b, he requests to be unlocked through the “left hand raise” gesture. In Figure 9c, the second participant requests to be locked via the “left hand waving” gesture. Then, the locked participant requests the UGV to follow him via the “left hand come” dynamic gesture (Figure 9d), while Figure 9e shows the UGV following this participant. Figure 9f depicts the participant, by raising his fist (“left hand fist”), ordering the UGV to pause the following mode. Figure 9g displays the participant loading a specially designed trolley with crates, as in this scenario, the UGV carries the crates from a random site of the orchard to a specific site. To that end, the movement of “right hand waving” (Figure 9h) is used for the UGV to start its autonomous navigation to the predefined location, as shown in Figure 8.

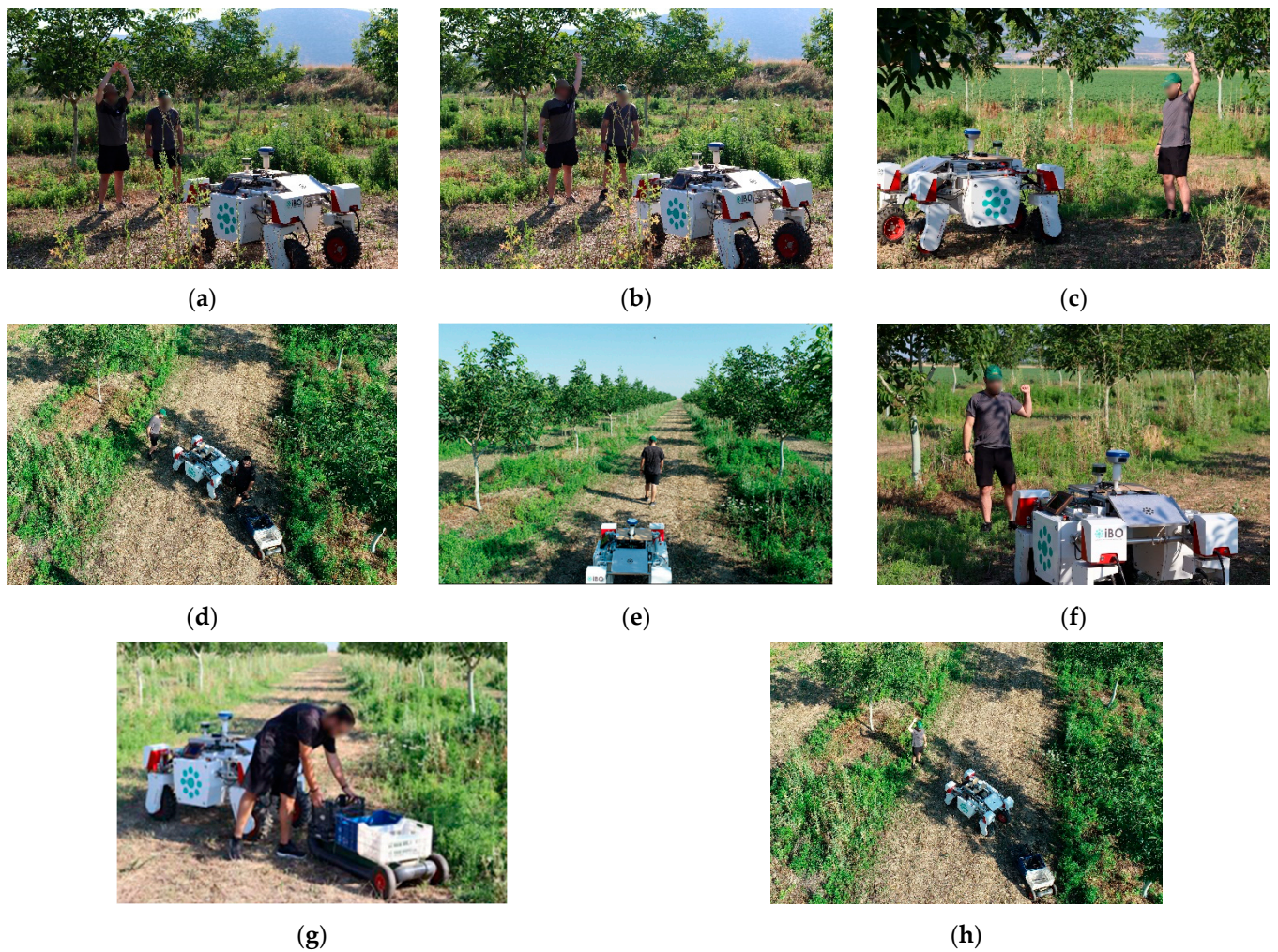


Figure 9. Images from the experimental demonstration of the integrated harvesting scenario in an orchard. In (a) the participant, who is currently locked, requests the UGV to stop all ongoing activities promptly; (b) the participant requests to be unlocked; (c) the second participant requests to be locked; (d) the participant signals the UGV to follow him; (e) the UGV follows the participant; (f) the participant signals the UGV to stop following him; (g) a participant fills a custom-built trolley with crates; (h) the participant indicates to the UGV to initiate its autonomous navigation to the preset destination.

4. Discussion and Conclusions

In this study, we focused on enhancing natural communication through a novel vision-based HRI by detecting and classifying a range of human movements. Various ML models were evaluated to determine their effectiveness in classifying these predefined human movements with LSTM demonstrating the highest overall performance. As a result, LSTM was chosen as the ML classifier for the proposed HRI framework.

Subsequently, the developed framework was rigorously tested through three carefully designed scenarios: (a) UGV following a participant; (b) GPS-based UGV navigation to a specific site of the orchard; and (c) an integrated harvesting scenario demonstrating all the developed HRI capabilities. In each scenario, safe limits for both the UGV's maximum speed and its minimum distance from participants and potential obstacles were thoroughly programmed and maintained.

In summary, all the designed scenarios were successfully tested, demonstrating the robustness and versatility of the developed HRI system in open-field agricultural settings.

The main challenges encountered and overcome during the implementation included the following:

- Human movement detection: In particular, the “left hand come” gesture was proved to be the most demanding, as the hand’s position in front of the body presents challenges in identifying specific reference landmarks accurately. However, the integration of whole-body detection in conjunction with the LSTM classifier significantly enhanced the system’s ability to interpret and respond to these dynamic movements effectively.
- Environmental conditions and background complexity: Ensuring reliable human movement recognition across different environmental conditions, such as varying lighting and background complexity, posed a major challenge. Toward ensuring adaptability, the system was successfully tested in multiple environments, including different sites of the orchard during both sunny and cloudy days.
- Navigation and obstacle avoidance: The UGV’s ability to navigate autonomously to predetermined locations required precise programming to avoid obstacles and always keep a safe speed and distance. For this purpose, the right selection of needed sensors in combination with the exploitation of the capabilities of ROS Navigation Stack enabled global localization and safe GPS-based navigation.

A notable distinction between the present study and the previous similar study [25] lies in the method of participant identification. While the previous study primarily relied on static gestures for detection, the present study introduced an innovative approach by utilizing dynamic human movements to trigger specific UGV actions. Incorporating dynamic gestures, this analysis demonstrated improved accuracy and responsiveness in interpreting human actions, enhancing the overall performance of the HRI framework. Indicatively, the best ML classifier in [25] was KNN with 0.839, 0.821, and 0.83 overall precision, recall and F1-score, respectively, while in this study, the corresponding values, provided by LSTM, were 0.912, 0.95, and 0.9, respectively. Thus, in contrast with the static hand gesture detection in [25], detecting reference landmarks from the whole body provided more accurate results under the same environmental conditions, which was mainly due to the larger body parts that should be identified. This shift from static to dynamic gestures represents also a significant advancement in enhancing the naturalness of human–robot communication.

Employing more natural ways of communication instills confidence, predictability and familiarity in the user, fostering “perceived safety” [39,48]. This sense of safety facilitates seamless collaboration between humans and robots, resulting in enhanced efficiency in agricultural operations [12]. To that end, innovative methods enabling real-time optimal control using output feedback [49] could improve the UGV’s ability to adapt to varying agricultural environments and tasks, making HRI more responsive to the immediate needs of human operators and environmental conditions.

Taking into account the recent progress in human activity recognition (HAR) through surface electromyography (sEMG) [50,51], this method could also be applied in agriculture, however, by carefully addressing their inherent drawbacks. The main limitations of pure sEMG-based HAR are its sensitivity to electrode positioning, various types of noise, challenges in interpreting deep learning models, and the requirement for extensive labeled datasets. Ovrur et al. [52] overcame these challenges by developing an autonomous learning framework that combines depth vision and electromyography to automatically classify EMG data, derived by a Myo armband device, based on depth information. Given also the ever-increasing use of smartphones in agriculture [53], the use of these devices could be considered for HAR. Indicatively, facing similar sEMG drawbacks, Qi et al. [54] addressed them by developing a skeleton-based framework utilizing IMU signals from a smartphone, a camera and depth sensors (Microsoft Kinect V2). Although different hand gestures were investigated in the aforementioned studies, the consistently high and stable performance observed suggests that these methods could be tested as an alternative to our purely vision-based framework in future investigations. However, any such evaluation must take into account the specific challenges posed by open-field environments.

In future research, further exploration could also focus on evaluating the system's performance in more complex and unpredictable environments with more people in the same area with the UGV. Interestingly, incorporating feedback from workers regarding their experiences and opinions of interacting with the UGV could provide valuable insights for refining the system and ensuring its practical usability. In addition, investigating the feasibility of allowing workers to operate at their own pace in real-world conditions could enhance the system's adaptability, trust and acceptance among users. In a nutshell, by considering both the efficiency, flexibility and safety of the system in tandem with the workers' perspectives, future research should endeavor to develop more socially accepted solutions for human–robot synergy in agricultural settings.

Author Contributions: Conceptualization, V.M. and D.B.; methodology, V.M., R.B., G.K. and E.P.; validation, A.P., L.B. and D.K.; formal analysis, R.B., E.P. and G.K.; investigation, V.M., D.K., A.P. and L.B.; writing—original draft preparation, V.M. and L.B.; writing—review and editing, E.P., G.K., D.K., R.B. and D.B.; visualization, V.M.; supervision, 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.

Data Availability Statement: The datasets presented in this article are not readily available because the data are part of an ongoing study. Requests to access the datasets should be directed to the authors of the article.

Conflicts of Interest: Author Vasileios Moysiadis and Dionysis Bochtis was employed by the company farmB Digital Agriculture S.A. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

1. Thakur, A.; Venu, S.; Gurusamy, M. An extensive review on agricultural robots with a focus on their perception systems. *Comput. Electron. Agric.* **2023**, *212*, 108146. [\[CrossRef\]](#)
2. Abbasi, R.; Martinez, P.; Ahmad, R. The digitization of agricultural industry—A systematic literature review on agriculture 4.0. *Smart Agric. Technol.* **2022**, *2*, 100042. [\[CrossRef\]](#)
3. Duong, L.N.K.; Al-Fadhli, M.; Jagtap, S.; Bader, F.; Martindale, W.; Swainson, M.; Paoli, A. A review of robotics and autonomous systems in the food industry: From the supply chains perspective. *Trends Food Sci. Technol.* **2020**, *106*, 355–364. [\[CrossRef\]](#)
4. Xie, D.; Chen, L.; Liu, L.; Chen, L.; Wang, H. Actuators and Sensors for Application in Agricultural Robots: A Review. *Machines* **2022**, *10*, 913. [\[CrossRef\]](#)
5. Bechar, A.; Vigneault, C. Agricultural robots for field operations: Concepts and components. *Biosyst. Eng.* **2016**, *149*, 94–111. [\[CrossRef\]](#)
6. Bechar, A.; Vigneault, C. Agricultural robots for field operations. Part 2: Operations and systems. *Biosyst. Eng.* **2017**, *153*, 110–128. [\[CrossRef\]](#)
7. Apraiz, A.; Lasa, G.; Mazmela, M. Evaluation of User Experience in Human–Robot Interaction: A Systematic Literature Review. *Int. J. Soc. Robot.* **2023**, *15*, 187–210. [\[CrossRef\]](#)
8. Abdulazeem, N.; Hu, Y. Human Factors Considerations for Quantifiable Human States in Physical Human–Robot Interaction: A Literature Review. *Sensors* **2023**, *23*, 7381. [\[CrossRef\]](#)
9. Semeraro, F.; Griffiths, A.; Cangelosi, A. Human–robot collaboration and machine learning: A systematic review of recent research. *Robot. Comput. Integr. Manuf.* **2023**, *79*, 102432. [\[CrossRef\]](#)
10. Tagarakis, A.C.; Benos, L.; Kyriakarakos, G.; Pearson, S.; Sørensen, C.G.; Bochtis, D. Digital Twins in Agriculture and Forestry: A Review. *Sensors* **2024**, *24*, 3117. [\[CrossRef\]](#)
11. Camarena, F.; Gonzalez-Mendoza, M.; Chang, L.; Cuevas-Ascencio, R. An Overview of the Vision-Based Human Action Recognition Field. *Math. Comput. Appl.* **2023**, *28*, 61. [\[CrossRef\]](#)
12. Benos, L.; Bechar, A.; Bochtis, D. Safety and ergonomics in human-robot interactive agricultural operations. *Biosyst. Eng.* **2020**, *200*, 55–72. [\[CrossRef\]](#)
13. Obrenovic, B.; Gu, X.; Wang, G.; Godinic, D.; Jakhongirov, I. Generative AI and human–robot interaction: Implications and future agenda for business, society and ethics. *AI Soc.* **2024**. [\[CrossRef\]](#)

14. Marinoudi, V.; Lampridi, M.; Kateris, D.; Pearson, S.; Sørensen, C.G.; Bochtis, D. The future of agricultural jobs in view of robotization. *Sustainability* **2021**, *13*, 12109. [CrossRef]
15. Benos, L.; Moysiadis, V.; Kateris, D.; Tagarakis, A.C.; Busato, P.; Pearson, S.; Bochtis, D. Human-Robot Interaction in Agriculture: A Systematic Review. *Sensors* **2023**, *23*, 6776. [CrossRef] [PubMed]
16. Aivazidou, E.; Tsolakis, N. Transitioning towards human–robot synergy in agriculture: A systems thinking perspective. *Syst. Res. Behav. Sci.* **2023**, *40*, 536–551. [CrossRef]
17. Adamides, G.; Edan, Y. Human–robot collaboration systems in agricultural tasks: A review and roadmap. *Comput. Electron. Agric.* **2023**, *204*, 107541. [CrossRef]
18. Lytridis, C.; Kaburlasos, V.G.; Pachidis, T.; Manios, M.; Vrochidou, E.; Kalampokas, T.; Chatzistamatis, S. An Overview of Cooperative Robotics in Agriculture. *Agronomy* **2021**, *11*, 1818. [CrossRef]
19. Yerebakan, M.O.; Hu, B. Human–Robot Collaboration in Modern Agriculture: A Review of the Current Research Landscape. *Adv. Intell. Syst.* **2024**, *6*, 2300823. [CrossRef]
20. Vasconez, J.P.; Kantor, G.A.; Auat Cheein, F.A. Human–robot interaction in agriculture: A survey and current challenges. *Biosyst. Eng.* **2019**, *179*, 35–48. [CrossRef]
21. Bonarini, A. Communication in Human-Robot Interaction. *Curr. Robot. Rep.* **2020**, *1*, 279–285. [CrossRef] [PubMed]
22. Robinson, F.; Pelikan, H.; Watanabe, K.; Damiano, L.; Bown, O.; Velonaki, M. Introduction to the Special Issue on Sound in Human-Robot Interaction. *J. Hum.-Robot Interact.* **2023**, *12*, 1–5. [CrossRef]
23. Ren, Q.; Hou, Y.; Botteldooren, D.; Belpaeme, T. No More Mumbles: Enhancing Robot Intelligibility Through Speech Adaptation. *IEEE Robot. Autom. Lett.* **2024**, *9*, 6162–6169. [CrossRef]
24. Wuth, J.; Correa, P.; Núñez, T.; Saavedra, M.; Yoma, N.B. The Role of Speech Technology in User Perception and Context Acquisition in HRI. *Int. J. Soc. Robot.* **2021**, *13*, 949–968. [CrossRef]
25. Moysiadis, V.; Katikaridis, D.; Benos, L.; Busato, P.; Anagnostis, A.; Kateris, D.; Pearson, S.; Bochtis, D. An Integrated Real-Time Hand Gesture Recognition Framework for Human-Robot Interaction in Agriculture. *Appl. Sci.* **2022**, *12*, 8160. [CrossRef]
26. Tagarakis, A.C.; Benos, L.; Aivazidou, E.; Anagnostis, A.; Kateris, D.; Bochtis, D. Wearable Sensors for Identifying Activity Signatures in Human-Robot Collaborative Agricultural Environments. *Eng. Proc.* **2021**, *9*, 9005. [CrossRef]
27. Tagarakis, A.C.; Filippou, E.; Kalaitzidis, D.; Benos, L.; Busato, P.; Bochtis, D. Proposing UGV and UAV Systems for 3D Mapping of Orchard Environments. *Sensors* **2022**, *22*, 1571. [CrossRef] [PubMed]
28. Niu, H.; Chen, Y. The Unmanned Ground Vehicles (UGVs) for Digital Agriculture. In *Smart Big Data in Digital Agriculture Applications: Acquisition, Advanced Analytics, and Plant Physiology-Informed Artificial Intelligence*; Niu, H., Chen, Y., Eds.; Springer: Cham, Switzerland, 2024; pp. 99–109. ISBN 978-3-031-52645-9.
29. Rondelli, V.; Franceschetti, B.; Mengoli, D. A Review of Current and Historical Research Contributions to the Development of Ground Autonomous Vehicles for Agriculture. *Sustainability* **2022**, *14*, 9221. [CrossRef]
30. Gokul, S.; Dhiksith, R.; Sundaresh, S.A.; Gopinath, M. Gesture Controlled Wireless Agricultural Weeding Robot. In Proceedings of the 2019 5th International Conference on Advanced Computing & Communication Systems (ICACCS), Coimbatore, India, 15–16 March 2019; pp. 926–929.
31. Yang, Z.; Jiang, D.; Sun, Y.; Tao, B.; Tong, X.; Jiang, G.; Xu, M.; Yun, J.; Liu, Y.; Chen, B.; et al. Dynamic Gesture Recognition Using Surface EMG Signals Based on Multi-Stream Residual Network. *Front. Bioeng. Biotechnol.* **2021**, *9*, 779353. [CrossRef]
32. Navigation: Package Summary. Available online: <https://www.opera.com/client/upgraded> (accessed on 28 June 2022).
33. Cao, Z.; Hidalgo, G.; Simon, T.; Wei, S.-E.; Sheikh, Y. OpenPose: Realtime Multi-Person 2D Pose Estimation using Part Affinity Fields. *arXiv* **2019**, arXiv:1812.08008. [CrossRef]
34. Dash, S.S.; Nayak, S.K.; Mishra, D. A review on machine learning algorithms. In Proceedings of the Smart Innovation, Systems and Technologies; Springer Science and Business Media, Faridabad, India, 14–16 February 2019; Deutschland GmbH: Berlin, Germany, 2021; Volume 153, pp. 495–507.
35. Arno, S.; Bell, C.P.; Alaia, M.J.; Singh, B.C.; Jazrawi, L.M.; Walker, P.S.; Bansal, A.; Garofolo, G.; Sherman, O.H. Does Anteromedial Portal Drilling Improve Footprint Placement in Anterior Cruciate Ligament Reconstruction? *Clin. Orthop. Relat. Res.* **2016**, *474*, 1679–1689. [CrossRef]
36. Labarrière, F.; Thomas, E.; Calistri, L.; Optasanu, V.; Gueugnon, M.; Ornetti, P.; Laroche, D. Machine Learning Approaches for Activity Recognition and/or Activity Prediction in Locomotion Assistive Devices—A Systematic Review. *Sensors* **2020**, *20*, 6345. [CrossRef] [PubMed]
37. Attri, I.; Awasthi, L.K.; Sharma, T.P. Machine learning in agriculture: A review of crop management applications. *Multimed. Tools Appl.* **2024**, *83*, 12875–12915. [CrossRef]
38. Pulver, A.; Lyu, S. LSTM with working memory. In Proceedings of the 2017 International Joint Conference on Neural Networks (IJCNN), Anchorage, AK, USA, 14–19 May 2017; pp. 845–851.
39. Benos, L.; Sørensen, C.G.; Bochtis, D. Field Deployment of Robotic Systems for Agriculture in Light of Key Safety, Labor, Ethics and Legislation Issues. *Curr. Robot. Rep.* **2022**, *3*, 49–56. [CrossRef]
40. Müller, M.; Ruppert, T.; Jazdi, N.; Weyrich, M. Self-improving situation awareness for human–robot-collaboration using intelligent Digital Twin. *J. Intell. Manuf.* **2024**, *35*, 2045–2063. [CrossRef]
41. Liu, Z.; Lü, Z.; Zheng, W.; Zhang, W.; Cheng, X. Design of obstacle avoidance controller for agricultural tractor based on ROS. *Int. J. Agric. Biol. Eng.* **2019**, *12*, 58–65. [CrossRef]

42. Gurevin, B.; Gulturk, F.; Yildiz, M.; Pehlivan, I.; Nguyen, T.T.; Caliskan, F.; Boru, B.; Yildiz, M.Z. A Novel GUI Design for Comparison of ROS-Based Mobile Robot Local Planners. *IEEE Access* **2023**, *11*, 125738–125748. [[CrossRef](#)]
43. Zheng, K. ROS Navigation Tuning Guide. In *Robot Operating System (ROS): The Complete Reference*; Koubaa, A., Ed.; Springer: Cham, Switzerland, 2021; Volume 6, pp. 197–226. ISBN 978-3-030-75472-3.
44. ROS 2 Documentation: Navfn. Available online: <http://wiki.ros.org/navfn> (accessed on 30 May 2024).
45. Cybulski, B.; Wegierska, A.; Granosik, G. Accuracy comparison of navigation local planners on ROS-based mobile robot. In Proceedings of the 2019 12th International Workshop on Robot Motion and Control (RoMoCo), Poznan, Poland, 8–10 July 2019; pp. 104–111.
46. Min, S.K.; Delgado, R.; Byoung, W.C. Comparative study of ROS on embedded system for a mobile robot. *J. Autom. Mob. Robot. Intell. Syst.* **2018**, *12*, 61–67.
47. Guo, L.; Lu, Z.; Yao, L. Human-Machine Interaction Sensing Technology Based on Hand Gesture Recognition: A Review. *IEEE Trans. Hum.-Mach. Syst.* **2021**, *51*, 300–309. [[CrossRef](#)]
48. Akalin, N.; Kristoffersson, A.; Loutfi, A. Do you feel safe with your robot? Factors influencing perceived safety in human-robot interaction based on subjective and objective measures. *Int. J. Hum. Comput. Stud.* **2022**, *158*, 102744. [[CrossRef](#)]
49. Zhao, J.; Lv, Y.; Zeng, Q.; Wan, L. Online Policy Learning-Based Output-Feedback Optimal Control of Continuous-Time Systems. *IEEE Trans. Circuits Syst. II Express Briefs* **2024**, *71*, 652–656. [[CrossRef](#)]
50. Rani, G.J.; Hashmi, M.F.; Gupta, A. Surface Electromyography and Artificial Intelligence for Human Activity Recognition—A Systematic Review on Methods, Emerging Trends Applications, Challenges, and Future Implementation. *IEEE Access* **2023**, *11*, 105140–105169. [[CrossRef](#)]
51. Mekruksavanich, S.; Jantawong, P.; Hnoohom, N.; Jitpattanakul, A. Human Activity Recognition for People with Knee Abnormality Using Surface Electromyography and Knee Angle Sensors. In Proceedings of the 2023 Joint International Conference on Digital Arts, Media and Technology with ECTI Northern Section Conference on Electrical, Electronics, Computer and Telecommunications Engineering (ECTI DAMT & NCON), Phuket, Thailand, 22–25 March 2023; pp. 483–487.
52. Ovrur, S.E.; Zhou, X.; Qi, W.; Zhang, L.; Hu, Y.; Su, H.; Ferrigno, G.; De Momi, E. A novel autonomous learning framework to enhance sEMG-based hand gesture recognition using depth information. *Biomed. Signal Process. Control* **2021**, *66*, 102444. [[CrossRef](#)]
53. Luo, X.; Zhu, S.; Song, Z. Quantifying the Income-Increasing Effect of Digital Agriculture: Take the New Agricultural Tools of Smartphone as an Example. *Int. J. Environ. Res. Public Health* **2023**, *20*, 3127. [[CrossRef](#)]
54. Qi, W.; Wang, N.; Su, H.; Aliverti, A. DCNN based human activity recognition framework with depth vision guiding. *Neurocomputing* **2022**, *486*, 261–271. [[CrossRef](#)]

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