

Benchmarking ML Approaches to UWB-Based Range-Only Posture Recognition for Human Robot-Interaction

Salma Salimi, Sahar Salimpour, Jorge Peña Queraltá, Wallace Moreira Bessa, and Tomi Westerlund

Abstract—Human pose estimation involves detecting and tracking the positions of various body parts using input data from sources such as images, videos, or motion and inertial sensors. This paper presents a novel approach to human pose estimation using machine learning algorithms to predict human posture and translate them into robot motion commands using ultra-wideband (UWB) nodes, as an alternative to motion sensors. The study utilizes five UWB sensors implemented on the human body to enable the classification of still poses and more robust posture recognition. This approach ensures effective posture recognition across a variety of subjects. These range measurements serve as input features for posture prediction models, which are implemented and compared for accuracy. For this purpose, machine learning algorithms including K-Nearest Neighbors (KNN), Support Vector Machine (SVM), and deep Multi-Layer Perceptron (MLP) neural network are employed and compared in predicting corresponding postures. We demonstrate the proposed approach for real-time control of different mobile/aerial robots with inference implemented in a ROS 2 node. Experimental results demonstrate the efficacy of the approach, showcasing successful prediction of human posture and corresponding robot movements with high accuracy.

Index Terms—Ultra-Wideband (UWB); Human Pose Estimation; K-Nearest Neighbors (KNN); Support Vector Machine (SVM); Multi-Layer Perceptron (MLP); ROS 2

I. INTRODUCTION

IN the current rapidly evolving technological landscape, robotics has assumed a pivotal role across diverse domains, from industrial automation and healthcare to entertainment and everyday tasks. As part of this integration, Human-Robot Interaction (HRI) facilitates the deployment of robots in human environments, spanning collaborative work settings such as manufacturing [1], service robots [2], search and rescue robots, and drones [3]. Robots with advanced sensors and AI, controlled by humans, can handle precise and strenuous tasks, while humans focus on decision-making.

Within HRI, the ability of robots to accurately perceive and interpret a wide range of human communication methods is essential [4]. Establishing a seamless communication system between humans and robots has emerged as a key research focus. When verbal commands or even neural signals are not the preferred mode of communication, human postures serve as an important channel for conveying intentions and commands [5]. Through body postures, individuals can convey discrete but precise signals to the robot to execute predefined tasks. Vision-based and wearable-based approaches dominate

this field when it comes to creating posture-based interaction interfaces [6].

Traditional human pose estimation methods predominantly relied on computer vision techniques, utilizing cameras such as RGB or infrared [7]. These systems face challenges like complicated backgrounds, occlusions, and low-resolution images. In recent years, there has been a significant trend towards 3D pose estimation. Accurate 3D pose annotations are often costly because they require specialized systems like motion capture with markers on the subject. Consequently, researchers have concentrated on creating techniques to derive 3D human pose data from 2D images. However, these 3D annotations are typically obtained in controlled environments, limiting their applicability in real-world scenarios. Some techniques have attempted to address this by using lifting operations to estimate 3D poses from 2D images captured by cameras [8]. Despite this, vision-based methods struggle with occlusions and depth ambiguity. For instance, transformer-based approaches have been used to mitigate partial occlusions by utilizing attention modules to capture global context and long-range relationships between the predicted joints [9]. Nonetheless, these methods still face challenges in scenarios with complete occlusion. In addition to vision-based systems, IMU sensors have been used to measure motion. However, IMU sensors can become inaccurate with movement, making it difficult to classify static postures. This inaccuracy further complicates the task of human posture estimation.

Recently, Ultra-Wide-Band (UWB) technology has emerged as a promising alternative due to its low complexity, high accuracy, and ability to penetrate obstacles [10]. UWB tech-

Salma Salimi, Faculty of Technology, University of Turku, Finland (e-mail: salmas@utu.fi).

Sahar Salimpour, Faculty of Technology, University of Turku, Finland (e-mail: sahar@utu.fi).

Jorge Peña Queraltá, Swiss Federal School of Technology, ETH Zürich, Switzerland (e-mail: jopequ@utu.fi).

Wallace Moreira Bessa, Faculty of Technology, University of Turku, Finland (e-mail: wmobes@utu.fi).

Tomi Westerlund, Faculty of Technology, University of Turku, Finland (e-mail: toveve@utu.fi).

nology offers a robust solution to the limitations of vision-based systems and IMUs. One implementation of UWB involves an impulse radio ultra-wideband (IR-UWB) radar system with an 8-by-8 multiple-input multiple-output (MIMO) antenna array, suitable for through-wall detection. The 3D-TransPOSE algorithm [7], based on transformer architecture, improves accuracy by focusing on relevant radar signal segments. However, it has limited penetration depth and complex hardware requirements. In contrast, methods using camera-equipped flying robots [11] offer real-time motion tracking but face challenges like high computational demands and safety concerns.

Over the past decade, UWB has evolved into a dependable, cost-effective, and commercially viable RF solution for data transmission, Time of Flight (TOF) or Time Difference of Arrival (TDOA) ranging, and localization [12]. Consequently, many roboticists have started integrating UWB into their projects due to its precision of approximately 10 centimeters, range of up to 100 meters, resilience to multipath interference, independence from line of sight (LOS), low power consumption, and communication speeds of 100 Mbit/s [13]. Despite these advantages, UWB measurements can still be affected by ranging errors and noise, which remains an active area of research in the robotics community [14]–[16].

In this paper, we propose utilizing UWB technology combined with transformer-based algorithms for human posture estimation. The micro-Doppler radar effect, for instance, has been employed to classify human movement without estimating body parts in [17], and a bistatic radar configuration using UWB has been developed to locate a moving person by analyzing the channel impulse response in [18]. Additionally, RF-Pose3D has shown the potential of using RF signals to infer 3D human poses, even with multiple individuals and obstacles such as walls and occlusions [19]. By leveraging these advances, our approach aims to overcome the limitations of traditional methods and provide more accurate and reliable human posture estimation.

We employ UWB sensors in our study for human posture estimation, a novel application to the best of our knowledge. By utilizing node distances placed on the human body, we aim to generate precise and reliable pose predictions. To validate our approach, we propose incorporating machine learning algorithms such as KNN, SVM, and MLP. Our model will be trained on UWB data to accurately predict human posture, allowing us to issue commands to robots based on these predictions.

Our approach offers several key contributions. Firstly, it provides enhanced accuracy, offering stable and accurate pose estimation regardless of motion. Secondly, it overcomes the limitations of traditional methods; unlike camera-based systems, our approach does not struggle with occlusions and depth ambiguity. Thirdly, UWB sensors are less demanding in terms of computational resources and power, making our approach more efficient. Finally, we provide a comprehensive dataset in our GitHub to support further research in the field, enhancing accessibility and reproducibility.

The remainder of this paper is organized as follows: Section II details the methodology and background on employing

different machine learning techniques with UWB technology for human pose estimation. Section III evaluates the classification performance of the various models introduced in Section II. Section IV presents the experimental results, and finally, Section V offers the conclusion.

II. METHODOLOGY

We performed human pose estimation using UWB sensor data by applying and evaluating various machine learning algorithms. The following sections delve into each step with comprehensive elaboration.

A. UWB Ranging

The high-accuracy short-range distance between two UWB nodes can be estimated using the ToF of a wireless signal exchanged between them, multiplied by the speed of light. ToF can be measured using one-way or two-way ranging, depending on device synchronization. One-way ranging requires precise clock synchronization as the signal travels in only one direction, whereas two-way ranging does not need synchronization, making it more commonly used due to its simplicity [20]. In this paper, we use Single-Sided Two-Way Ranging (SS-TWR), where the distance is calculated based on the time of flight of a UWB signal query and its reply. Errors in range estimates between UWB nodes, which depend on different environmental error sources, LOS or Non-Line of Sight (NLOS) conditions, can generally be modeled as Gaussian random variables [21]. In optimal LOS scenarios, the error can be as low as 20 centimeters [21]. Additionally, we use Decawave DWM1001 UWB modules, which offer accuracy of up to 5 centimeters. Data is collected using the ROS 2 with custom packages developed to interface with the modules.

B. Data Acquisition

Five volunteers actively participated in this study, with UWB nodes strategically placed on their ankles, wrists, and bellies to comprehensively capture a range of movements and postures. Engaging in nine distinct postures—standby, up, down, left, right, takeoff, land, forward, and backward—these individuals facilitated the collection of rich UWB data, providing a robust foundation for our model to accurately predict corresponding poses, which is particularly pertinent for guiding robots. The dataset presented in this work consists of these nine different human postures, utilizing Qorvo's DWM1001 UWB modules with custom firmware for ToF ranging, all placed on the participants' bodies.

All measurements are obtained from the DWM1001 device via UART connection and interfaced with ROS 2 nodes for processing and recording. The five UWB nodes, positioned on the human body, communicate in an all-to-all manner, taking UWB measurements of the distances between nodes at a frequency of 15 Hz. For each posture, 400 measurements are recorded, ensuring a comprehensive dataset that captures the human movement and provides reliable data for model training and validation.

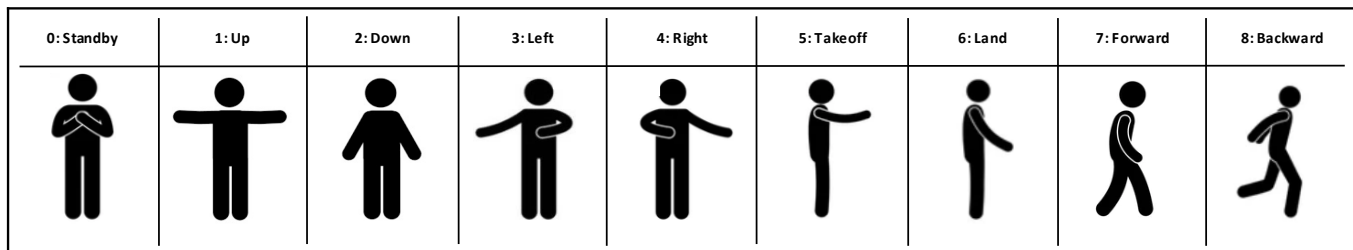


Fig. 1. Showcase of the different classes.

C. Machine Learning algorithms

Machine Learning aims to enable systems to learn from data, making predictions or identifying patterns [22]. Machine Learning methods are generally classified into supervised, unsupervised, and reinforcement learning [22], [23] where reinforcement learning optimizes actions through iterative interactions with the environment, guided by a reward system [24].

In this research, we collected UWB data from various human postures at a frequency of 15 Hz and employed machine learning and deep learning algorithms for human gesture prediction. Specifically, we compared the performance and accuracy of three classifiers: KNN, SVM, and MLP [25], [26], implemented using the scikit-learn (sklearn) library. The aim was to evaluate the efficiency of each method in accurately predicting human gestures and to identify which algorithm performs best under different conditions. This comparative analysis helps highlight the strengths and limitations of each approach within the proposed UWB-based posture estimation system, providing insights into their suitability for real-world applications. By understanding the performance variations among these models, we can better select the most appropriate algorithm for specific tasks, optimize computational resources, and improve the overall robustness and reliability of the human pose estimation system.

1) *KNN*: The KNN algorithm [27] classifies data points based on the proximity of k -nearest neighbors, using the Manhattan distance for evaluation. Classification is determined by the majority label among these neighbors. KNN has been widely used in fingerprint identification techniques [28]–[30], particularly for estimating the position of target devices. However, it requires maintaining large radio maps for frequent computations.

2) *SVM*: The SVM, a classical algorithm [31], identifies the optimal hyperplane to separate data into classes, maximizing the margin between them. It is widely used in classification tasks, modeling both linear and non-linear relationships [32], [33]. The kernel mechanism enhances its generalization, making SVM suitable for various fingerprint identification applications.

3) *MLP*: The MLP architecture, an artificial neural network [34], consists of an input layer, multiple hidden layers, and an output layer, with each layer containing neuron models. Neurons in adjacent layers are fully interconnected, enabling computation propagation from input to output during inference. Supervised training is performed using the back-propagation method, allowing MLP to transform input data

into target results effectively. MLP's robust feature extraction capabilities make it suitable for various applications in neural networks and fingerprint analysis [35].

III. POSE CLASSIFICATION

In this section, we present the classification performance of the different models introduced in the previous section. Participants for this study were selected randomly, without specific group criteria, to reflect a range of body types and ensure diversity in the dataset. A total of five individuals were chosen, each with distinct heights and weights, allowing for a variety of body shapes, which is crucial for testing the robustness of the models in recognizing different postures. For each posture, 400 measurements were recorded from UWB nodes for these participants. The data from four participants was used to train our models, while the data from the remaining participant was used to test the models. This approach ensures that the models are evaluated on unseen data, objectively assessing their predictive capabilities. For our posture prediction, we defined nine distinct classes from 0 to 8, as shown in Fig. 1. This classification framework allows for precise identification and prediction of various human postures, which is essential for guiding robotic movements.

To further evaluate the performance and robustness of our models, we introduced noise ranging from 0 to 30 centimeters separately to the training data, the test data, or both. This testing scenario helps determine how well the models can handle real-world data imperfections and maintain accuracy under varying conditions. For data collection, we exclusively used UWB nodes, with the tags affected only by materials that obstruct them, resulting in NLOS noise. To account for this noise in our model training, we gathered data from various individuals, as the only obstacle between the tags was the human body, dressed in different types of clothing.

For validating the models, we use a leave-one-out cross-validation methodology. For each iteration, the data from four participants was used for training, and the model was tested on the unseen data from the fifth participant, ensuring that it was evaluated on completely new data. With this, we aim at demonstrating generalization as the models are not biased towards any particular individual's data while avoiding overfitting to the relatively limited amount of data, from the perspective of having different test subjects.

By thoroughly testing the models under these various conditions, we aim to identify the most robust and accurate approach

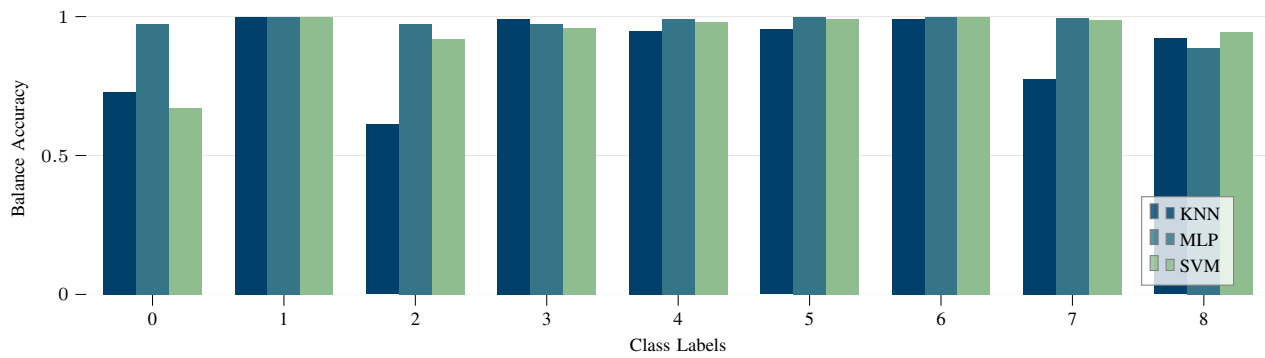


Fig. 2. Each class balanced accuracy.

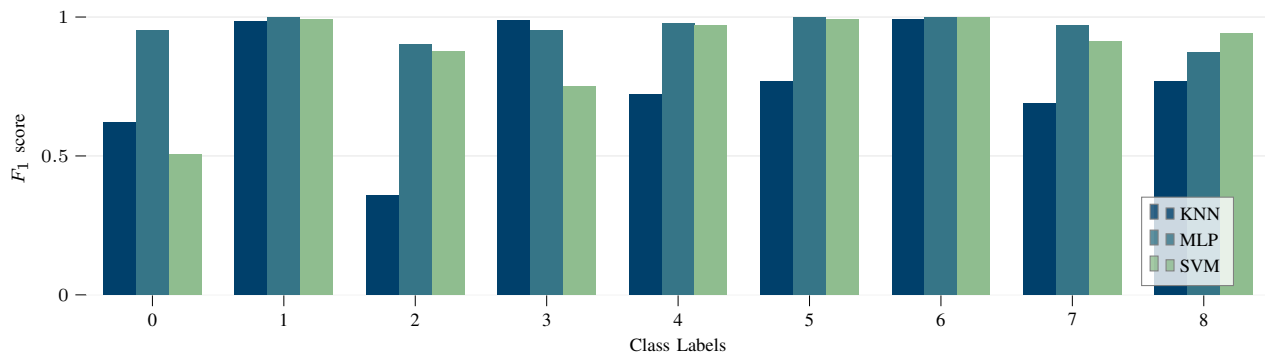


Fig. 3. Each class F_1 score.

for human gesture prediction, ensuring that the chosen model can perform effectively in diverse real-world scenarios.

1) *KNN*: For the KNN model, the optimal value of k was determined using the elbow method. In the absence of noise, the model achieved an accuracy of 86% for the test participant's data with the best k value being 2 where the confusion matrix can be seen in Fig. 4. Detailed balanced accuracy and F_1 score for each class are illustrated in Fig. 2 and Fig. 3, respectively. As shown, the model achieved a balanced accuracy of 73% for class 0, indicating moderate performance. However, the model struggled with class 2, achieving only a 61% balanced accuracy. These results are mirrored in the F_1 scores, where class 0 has an F_1 score of 62%, and class 2 shows an even lower F_1 score of 36%. The model performed exceptionally well for other classes.

When noise was introduced, ranging from 0 to 30 centimeters, the model's performance varied. With noise added to the training data, the accuracy dropped to 79%, and the optimal k value increased to 4. Adding noise solely to the test data resulted in an accuracy of 75%, while keeping the best k value at 2. Introducing noise to both training and test data led to an accuracy of 76%, with the optimal k value remaining at 2. These results indicate that while KNN is sensitive to noise, it retains reasonable performance with appropriate adjustments to the k value across different scenarios.

2) *SVM*: For the SVM model, Gridsearch was used to find the optimal parameters, such as C and γ values. In the absence of noise, the model achieved an accuracy of 97% with C equal to 1 and γ equal to 0.1 where the confusion matrix can be seen in Fig. 5. As it can be seen in

0	183	0	0	0	0	217	0	0	0
1	0	400	0	0	0	0	0	0	0
2	0	0	92	0	303	2	0	3	0
3	1	0	0	392	0	7	0	0	0
4	0	0	0	0	396	0	0	4	0
5	0	6	3	0	0	391	0	0	0
6	5	0	0	1	0	0	394	0	0
7	0	0	0	0	0	0	0	223	177
8	0	6	18	0	0	0	0	17	359
	0	1	2	3	4	5	6	7	8
	0	1	2	3	4	5	6	7	8

Fig. 4. KNN confusion matrix obtained with leave-one-out cross-validation (iterating over all subjects).

Fig. 2, SVM model exhibited robust performance across all classes, maintaining high balanced accuracy and F_1 scores. The model achieved a perfect balanced accuracy of 100% for class 6 and 1, reflecting flawless performance, and maintained performance above 95% for most other classes. The F_1 scores as shown in Fig. 3 corroborate these findings, with a perfect score for classes 6 and 2 and high scores for other classes, including 97% for class 4 and 99% for class 5.

When noise levels between 0 and 30 centimeters were applied, the model's accuracy decreased to 86% when noise



Fig. 5. SVM confusion matrix obtained with leave-one-out cross-validation (iterating over all subjects).

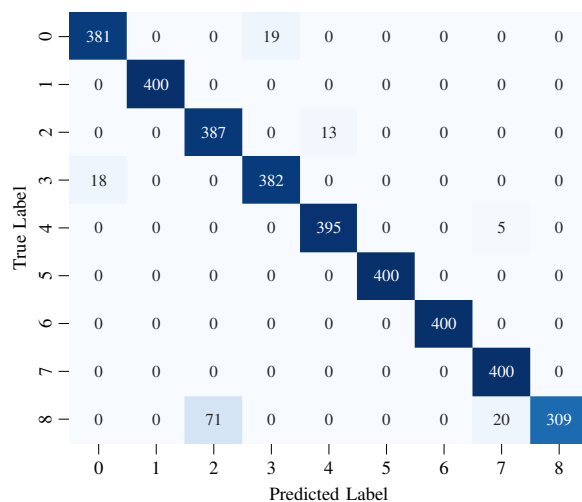


Fig. 6. MLP confusion matrix obtained with leave-one-out cross-validation (iterating over all subjects).

was added to the training data, with the optimal C value shifting to 10 while gamma remained at 0.1. Introducing noise only in the test data reduced the accuracy to 81%, maintaining C at 1 and gamma at 0.1. When noise was present in both the training and test data, the accuracy improved to 92%, with C set to 10 and gamma to 0.1. These results underscore the SVM model’s robustness and adaptability to noise, maintaining high accuracy across varying conditions.

3) *MLP*: For the MLP model, the accuracy was 96% in the absence of noise where the confusion matrix can be seen in Fig. 6. As can be seen in Fig. 2 and Fig. 3, MLP model demonstrated exceptional performance in human posture estimation, achieving high balanced accuracy and F1 scores across most classes. The model achieved perfect balanced accuracy for classes 1, 5, 6 and 7, reflecting its strong classification capabilities. Corresponding F1 scores were also perfect for classes 1,5 and 6, with additional high scores of 98% for class 4 and 97% for class 7. While the model showed slight performance dips for classes 2 and 8, the balanced accuracies obtained were 97% and 89%, respectively, and F1 scores of 90% and 87%.

With the introduction of noise between 0 and 30 centimeters, the model’s accuracy dropped to 83% when noise was added to the training data. Noise only in the test data reduced accuracy to 80%. Finally, when noise was introduced to both training and test data, the accuracy improved to 89%. These results highlight the MLP model’s robustness, showing it can maintain strong performance even in the presence of noise, making it a reliable choice for precise human posture estimation tasks.

In this study, we used five UWB nodes for data recording and model training to predict human postures accurately and send movement commands to robots. To determine the minimum number of nodes required for accurate posture prediction, we tested the same procedure with different numbers of UWB nodes, as the average balanced accuracy illustrated in Fig. 7. The results demonstrate a clear trend: the reduction in the number of nodes generally leads to a decrease in

balanced accuracy across all models. With five nodes, the MLP model achieved the highest balanced accuracy at 0.89, followed by SVM at 0.96 and KNN at 0.79. As the number of nodes decreased to four, all models experienced a slight drop in performance, with MLP and SVM still maintaining high balanced accuracy values of 0.84 and 0.85, respectively, while KNN dropped to 0.76. Further reductions to three nodes resulted in a more noticeable decline in accuracy for all models indicating that three nodes might be nearing the lower limit for effective posture prediction. When reduced to two nodes, the performance drop was significant suggesting that two nodes are insufficient for reliable posture estimation.

These findings highlight the importance of the number of UWB nodes in maintaining high prediction accuracy. While the MLP and SVM models showed some resilience to the reduction in nodes, the overall trend indicates that a minimum of four nodes is necessary to achieve reasonably accurate posture predictions. This reduction in node count can streamline the setup and simplify deployment while still ensuring reliable performance for robotic movement commands.

IV. EXPERIMENTAL RESULTS

In this section, we analyze the performance of the SVM model presented in the previous section, as it exhibited the highest performance among the tested models. This evaluation aims to demonstrate the robustness and practicality of our approach in real-world scenarios. By focusing on the SVM model, we can highlight the challenges and limitations that arise in practical applications. To provide a comprehensive assessment, we conducted experiments in two different real-world scenarios. The first scenario involved using one single drone, while the second scenario included both drone and ground robot. Implementing the SVM model in these varied setups allowed us to observe its effectiveness and identify potential areas for improvement. For our posture prediction, we defined nine distinct classes for the drone: 0 for standby, 1 for up, 2 for down, 3 for takeoff, 4 for land, 5 for left, 6 for



Fig. 7. Assessment of the maximum and minimum counts of UWB nodes.

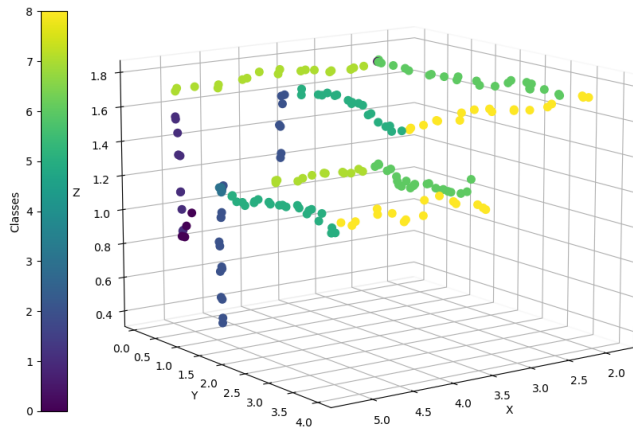


Fig. 8. Drone 3D path trajectory.

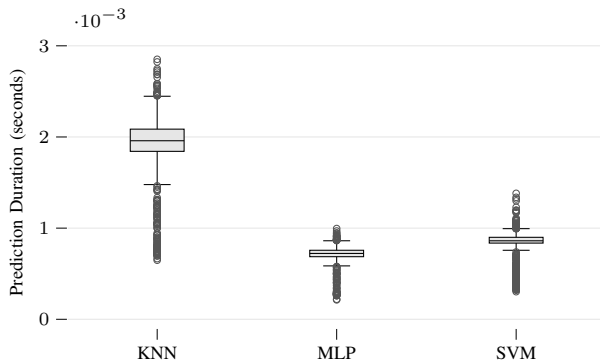


Fig. 9. Model prediction duration throughout the real-world experiment.

right, 7 for forward, and 8 for backward movement. For the ground robot, we used only classes 0, 5, 6, 7, and 8.

A. Experimental Setup

The experimental platform for this study comprises a commercially available Ryze Tello MAV and Turtlebot4 platform both equipped with OptiTrack motion capture (MOCAP) system markers. This setup enables precise recording of the robot's trajectory while executing movement commands, allowing us to monitor their movements and verify the accuracy of our model predictions.

For ground truth, we utilized a MOCAP system consisting of six OptiTrack cameras connected to a backbone WiFi network. This system covers an operating area approximately 8 meters wide, 9 meters long, and 5 meters high.

In this experiment, a separate individual, distinct from the five participants who provided training and testing data for the model, was equipped with five DWM1001-Dev UWB

nodes programmed with custom firmware for measuring distances [36]. Five of these nodes were strategically placed on the belly, wrists, and ankles to ensure comprehensive data collection. Additionally, two UWB nodes were connected to a laptop: one node functioned as a listener, while the other served as the network controller. The network controller node is responsible for resetting and initiating the range calculation when needed and is controlled by a ROS 2 node running on the laptop. All ROS 2 nodes responsible for processing and recording the data were run on this laptop. This configuration allowed us to efficiently collect distance measurements from the UWB nodes, which were then processed and analyzed within the ROS 2 framework. Therefore, the measurements obtained from the DWM1001 devices were indeed managed through wired connections to the laptop.

Furthermore, in the real-world experiments, we selected a participant who was not part of the training dataset, ensuring that the model's performance was evaluated on entirely new, unseen data. This approach highlights the model's generalizability across different body types, reinforcing its robustness in various practical applications.

The primary system for training the models and making predictions operates on Ubuntu 20.04. All nodes responsible for receiving UWB data, estimating postures, and sending movement commands to the robots were running ROS 2 Galactic. All the recorded data from our experiments are available in the project's repository¹.

B. Obtained Results

To demonstrate the versatility and capability of mapping a wide array of human postures or classes to either specific degrees of freedom in robots or to more actions in complex robots and multi-robot systems, we conducted two real-world experiments. The first experiment involved controlling only a Tello drone, while the second involved a Tello drone and a TurtleBot4 high-level teleoperation. In both experiments, a SVM model was used for human pose estimation.

As shown in Fig. 8, during the experiment utilizing the SVM model for posture estimation, the participant successfully maneuvered the drone within two designated squares at varying heights. The image illustrates the 3D movement of the drone, highlighting its ability to navigate through space accurately. Additionally, the 3D plot includes a color bar indicating the posture predicted at each movement, demonstrating that the predictions were quite accurate and allowed the participant to control the drone with ease. This successful demonstration

¹<https://github.com/salmasalimii/UWB-based-posture-recognition>

underscores the robustness and practical applicability of our system in real-world scenarios. The ability to accurately interpret human postures and translate them into precise drone movements underscores the potential of our approach for various applications in human-robot interaction and autonomous systems.

Furthermore, in Fig. 9, we calculated the prediction duration during our experiment to assess the latency in prediction calculation across all three models. For the KNN model, the prediction duration was around 0.002 seconds on average, while for SVM and MLP, it was less than 0.001 seconds. This analysis highlights the efficiency of our models, particularly the SVM model, in providing timely and accurate posture predictions, which is crucial for real-time control in autonomous systems. Additionally, the accuracy of all three models during the real-world experiment is summarized in Table I, providing a direct comparison of their real-world performance.

ML algorithms	SVM	MLP	KNN
real world accuracy	96%	94%	89%

TABLE I

ACCURACY OF ALL ALGORITHMS DURING THE REAL-WORLD EXPERIMENT

In the second scenario, we employed two robots: a Tello drone and a TurtleBot. The goal of this experiment was to control both robots while attempting to follow a square pattern in the experimental arena. Fig. 10 illustrates the actual UWB ranges collected during the experiment. In our system, we implemented a filter that discards distance measurements exceeding 2.5 meters, maintaining the previous classification as the current predicted class. The posture predictions made by the SVM model are also shown in the figure, along with the linear and angular velocity commands sent to the robots for clearer visualization. During the experiment, the participant controlled the robots in the square pattern while adjusting the drone's altitude. As the TurtleBot's height is fixed, specific postures were designated solely for commanding the drone's vertical movements. exclusively to command the drone's vertical movements.

As it can be seen in Fig 10 the SVM model maintained sufficient accuracy to effectively control the robots. This demonstrates the robustness of our approach in handling real-world noise and ensuring reliable performance. Fig. 11 depicts the trajectory tracking of both robots, indicating that our model accurately predicted the postures, allowing the participant to guide the robots as intended. The successful coordination between the drone and the TurtleBot highlights the model's capability to facilitate precise and responsive control in multi-robot systems, underscoring its potential for practical applications in complex environments.

V. DISCUSSION

In this study, we have conducted a comprehensive comparison of the KNN, SVM, and MLP models, particularly evaluating their latency and performance in posture prediction

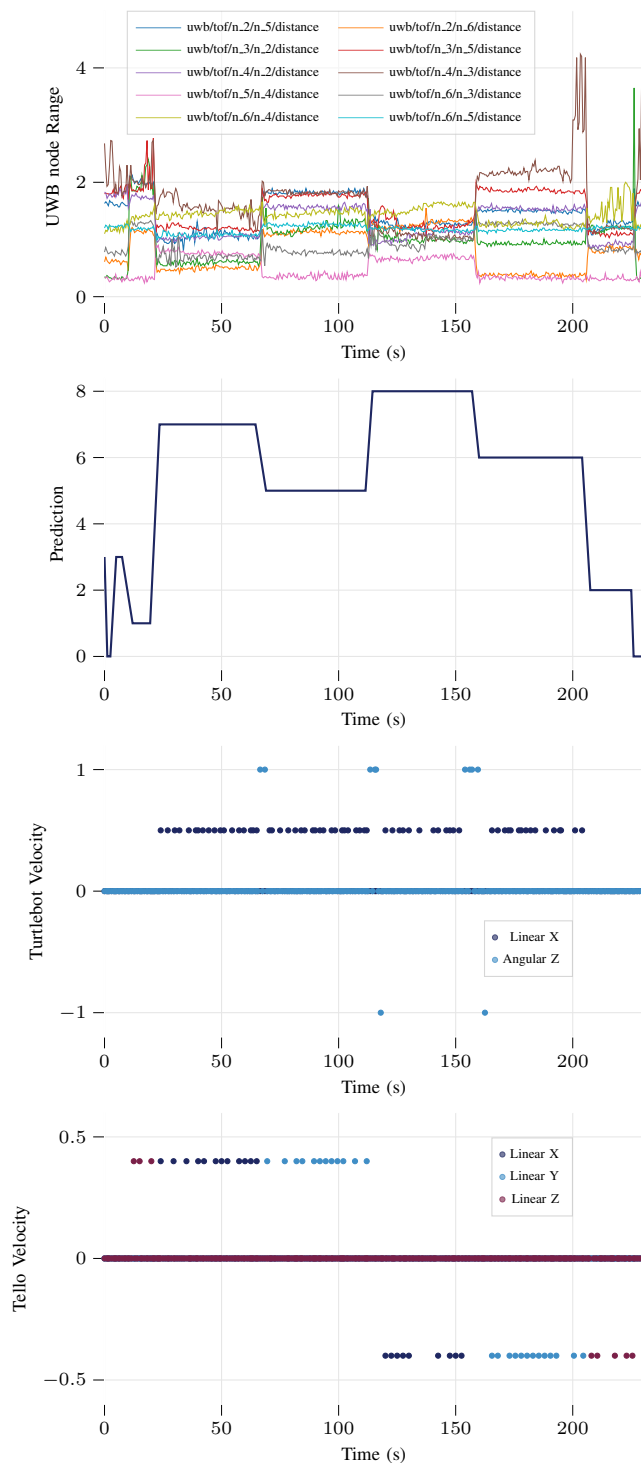


Fig. 10. Second scenario experiment results.

tasks. The analysis highlights the strengths and weaknesses of each model in different contexts such as performance under noisy conditions, computational latency, and scalability for real-time applications. Furthermore, two real-world experiments were carried out using the SVM model to further assess its effectiveness. The results of these experiments demonstrate the practical applicability of our approach, offering valuable insights into its performance under real-world conditions. This analysis provides a basis for understanding the trade-offs and

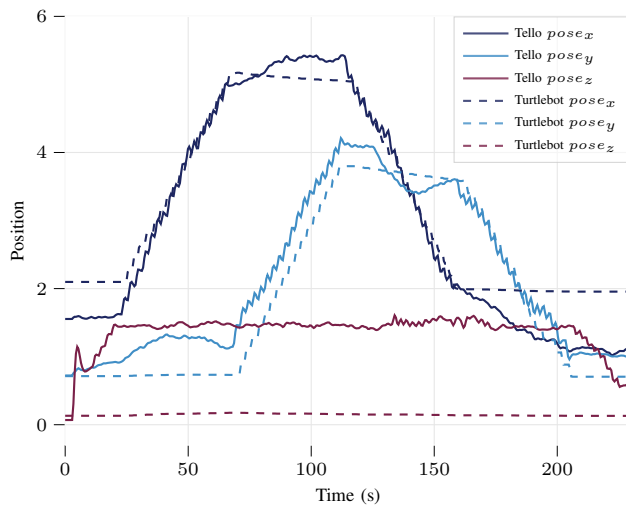


Fig. 11. Tello and Turtlebot trajectory.

identifying potential areas for further refinement.

While the current study focuses on single-person posture prediction, it also opens avenues for future work in more complex scenarios, such as multi-person pose tracking. In such cases, the system would likely require additional UWB nodes for each individual, necessitating retraining to accommodate the increased complexity of the data. Implementing multi-person tracking in dynamic environments introduces additional challenges, as obstacles between individuals or tags may result in NLOS noise or multipath issues in larger settings. To effectively address these challenges, it will be essential to separately account for NLOS and LOS conditions during the training process.

VI. CONCLUSION

We have presented an evaluation of three machine learning models—KNN, SVM, and MLP—for human posture estimation. Our study included a detailed real-world examination of SVM's performance, which generally showed higher accuracy compared to MLP and KNN. Across real-world experiments involving two scenarios with multi-robots, SVM's performance was analyzed, highlighting its competitive accuracy and efficiency in predicting human postures. In both controlled settings and real-world applications, our approach demonstrated practical utility and responsiveness, showing variations in performance across different posture classes and in the presence of noise. Despite these challenges, our approach's adaptability to noise and environmental factors underscores its robustness, making it suitable for applications requiring precise human-robot interaction and autonomous system control.

REFERENCES

- [1] Shitij Kumar, Celal Savur, and Ferat Sahin. Survey of human-robot collaboration in industrial settings: Awareness, intelligence, and compliance. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 51(1):280–297, 2020.
- [2] Abolfazl Mohebbi. Human-robot interaction in rehabilitation and assistance: a review. *Current Robotics Reports*, 1(3):131–144, 2020.
- [3] Mingyang Lyu, Yibo Zhao, Chao Huang, and Hailong Huang. Unmanned aerial vehicles for search and rescue: A survey. *Remote Sensing*, 15(13):3266, 2023.
- [4] Hang Su, Wen Qi, Jiahao Chen, Chenguang Yang, Juan Sandoval, and Med Amine Laribi. Recent advancements in multimodal human-robot interaction. *Frontiers in Neurorobotics*, 17:1084000, 2023.
- [5] Pedro Neto, Miguel Simão, Nuno Mendes, and Mohammad Safeea. Gesture-based human-robot interaction for human assistance in manufacturing. *The International Journal of Advanced Manufacturing Technology*, 101:119–135, 2019.
- [6] Valeria Villani, Cristian Secchi, Marco Lippi, and Lorenzo Sabatini. A general pipeline for online gesture recognition in human-robot interaction. *IEEE Transactions on Human-Machine Systems*, 53(2):315–324, 2023.
- [7] Gon Woo Kim, Sang Won Lee, Ha Young Son, and Kae Won Choi. A study on 3d human pose estimation using through-wall ir-uwb radar and transformer. *IEEE Access*, 11:15082–15095, 2023.
- [8] Adam Gosztołai, Semih Günel, Victor Lobato Ríos, Marco Pietro Abrate, Daniel Morales, Helge Rhodin, Pascal Fua, and Pavan Ramdya. Liftpose3d, a deep learning-based approach for transforming 2d to 3d pose in laboratory animals. *BioRxiv*, pages 2020–09, 2020.
- [9] Alexey Dosovitskiy, Lucas Beyer, Alexander Kolesnikov, Dirk Weissenborn, Xiaohua Zhai, Thomas Unterthiner, Mostafa Dehghani, Matthias Minderer, Georg Heigold, Sylvain Gelly, et al. An image is worth 16x16 words: Transformers for image recognition at scale. *arXiv preprint arXiv:2010.11929*, 2020.
- [10] Giulia Martinelli, Luca Santoro, Matteo Nardello, Davide Brunelli, Daniele Fontanelli, and Nicola Conci. Unposed: an ultra-wideband network for pose estimation with deep learning. In *2023 IEEE International Workshop on Metrology for Industry 4.0 & IoT (MetroInd4. 0&IoT)*, pages 299–304. IEEE, 2023.
- [11] Tobias Nägeli, Samuel Oberholzer, Silvan Plüss, Javier Alonso-Mora, and Otmar Hilliges. Flycon: Real-time environment-independent multi-view human pose estimation with aerial vehicles. *ACM Transactions on Graphics (TOG)*, 37(6):1–14, 2018.
- [12] Andrew Fishberg and Jonathan P How. Multi-agent relative pose estimation with uwb and constrained communications. In *2022 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 778–785. IEEE, 2022.
- [13] Abdulrahman Alarifi, AbdulMalik Al-Salman, Mansour Alsaleh, Ahmad Alnafessah, Suheer Al-Hadhrami, Mai A Al-Ammar, and Hend S Al-Khalifa. Ultra wideband indoor positioning technologies: Analysis and recent advances. *Sensors*, 16(5):707, 2016.
- [14] Nour Smaoui, Omprakash Gnawali, and Kyungki Kim. Study and mitigation of platform related uwb ranging errors. In *2020 International Conference on COMMunication Systems & NETWORKS (COMSNETS)*, pages 346–353. IEEE, 2020.
- [15] Michael Hamer and Raffaello D'Andrea. Self-calibrating ultra-wideband network supporting multi-robot localization. *Ieee Access*, 6:22292–22304, 2018.
- [16] Anton Ledergerber and Raffaello D'andrea. Calibrating away inaccuracies in ultra wideband range measurements: A maximum likelihood approach. *IEEE Access*, 6:78719–78730, 2018.
- [17] Hao Du, Tian Jin, Yongping Song, Yongpeng Dai, and Meng Li. A three-dimensional deep learning framework for human behavior analysis using range-doppler time points. *IEEE geoscience and remote sensing letters*, 17(4):611–615, 2019.
- [18] Maria Doglioni, Luca Santoro, Matteo Nardello, Daniele Fontanelli, and Davide Brunelli. Cost-effective bistatic radar with ultrawide-band radio. In *2022 IEEE International Workshop on Metrology for Industry 4.0 & IoT (MetroInd4. 0&IoT)*, pages 207–211. IEEE, 2022.
- [19] Mingmin Zhao, Yonglong Tian, Hang Zhao, Mohammad Abu Alsheikh, Tianhong Li, Rumen Hristov, Zachary Kabelac, Dina Katabi, and Antonio Torralba. Rf-based 3d skeletons. In *Proceedings of the 2018 Conference of the ACM Special Interest Group on Data Communication*, pages 267–281, 2018.
- [20] Jorge Peña Queraltá, Carmen Martínez Almansa, Fabrizio Schiano, Dario Floreano, and Tomi Westerlund. Uwb-based system for uav localization in gnss-denied environments: Characterization and dataset. In *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 4521–4528. IEEE, 2020.
- [21] Mengyao Dong. A low-cost nlos identification and mitigation method for uwb ranging in static and dynamic environments. *IEEE Communications Letters*, 25(7):2420–2424, 2021.
- [22] Di Wu, Renchi Zhang, Ameya Pore, Diego Dall'Alba, Xuan Thao Ha, Zhen Li, Yao Zhang, Fernando Herrera, Mouloud Ourak, Wojtek Kowalczyk, et al. A review on machine learning in flexible surgical and interventional robots: Where we are and where we are going. *Biomedical Signal Processing and Control*, 93:106179, 2024.
- [23] Ethem Alpaydin. *Introduction to machine learning*. MIT press, 2020.

- [24] Richard S Sutton and Andrew G Barto. *Reinforcement learning: An introduction*. MIT press, 2018.
- [25] Adam Krawczyk, Jakub Marciniak, and Dominik Belter. Comparison of machine learning techniques for self-collisions checking of manipulating robots. In *2023 27th International Conference on Methods and Models in Automation and Robotics (MMAR)*, pages 472–477. IEEE, 2023.
- [26] Xuxin Lin, Jianwen Gan, Chaohao Jiang, Shuai Xue, and Yanyan Liang. Wi-fi-based indoor localization and navigation: A robot-aided hybrid deep learning approach. *Sensors*, 23(14):6320, 2023.
- [27] Oliver Kramer et al. *Dimensionality reduction with unsupervised nearest neighbors*, volume 51. Springer, 2013.
- [28] Feng Yu, Minghua Jiang, Jing Liang, Xiao Qin, Ming Hu, Tao Peng, and Xinrong Hu. 5 g wifi signal-based indoor localization system using cluster k-nearest neighbor algorithm. *International journal of distributed sensor networks*, 10(12):247525, 2014.
- [29] Yaqin Xie, Yan Wang, Arumugam Nallanathan, and Lina Wang. An improved k-nearest-neighbor indoor localization method based on spearman distance. *IEEE signal processing letters*, 23(3):351–355, 2016.
- [30] Minh Tu Hoang, Yizhou Zhu, Brosnan Yuen, Tyler Reese, Xiaodai Dong, Tao Lu, Robert Westendorp, and Michael Xie. A soft range limited k-nearest neighbors algorithm for indoor localization enhancement. *IEEE Sensors Journal*, 18(24):10208–10216, 2018.
- [31] Shan Suthaharan. Machine learning models and algorithms for big data classification. *Integr. Ser. Inf. Syst*, 36:1–12, 2016.
- [32] Kadir Sabanci, Enes Yigit, Deniz Ustun, Abdurrahim Toktas, and Muhammet Fatih Aslan. Wifi based indoor localization: application and comparison of machine learning algorithms. In *2018 XXIIIrd International Seminar/Workshop on Direct and Inverse Problems of Electromagnetic and Acoustic Wave Theory (DIPED)*, pages 246–251. IEEE, 2018.
- [33] Feng Yu, Ming Hua Jiang, Jing Liang, Xiao Qin, Ming Hu, Tao Peng, and Xin Rong Hu. An indoor localization of wifi based on support vector machines. *Advanced Materials Research*, 926:2438–2441, 2014.
- [34] M Paegelow and LUC Budget. In geomatic approaches for modeling land change scenarios. *Camacho Omedo, MT, Paegelow, M., Mas, JF, Escobar, F., Eds*, pages 437–440, 2018.
- [35] Qiwu Zhu, Qingyu Xiong, Kai Wang, Wang Lu, and Tong Liu. Accurate wifi-based indoor localization by using fuzzy classifier and mlps ensemble in complex environment. *Journal of the Franklin Institute*, 357(3):1420–1436, 2020.
- [36] Paola Torrico Morón, Jorge Peña Queralta, and Tomi Westerlund. Towards large-scale relative localization in multi-robot systems with dynamic uwb role allocation. In *2022 7th International Conference on Robotics and Automation Engineering (ICRAE)*, pages 239–246. IEEE, 2022.



autonomy, distributed perception, and robot learning.

Jorge Peña Queralta is a postdoctoral researcher at Swiss Federal School of Technology, ETH, Switzerland. He received B.Sc. degrees in mathematics and physics engineering from UPC BarcelonaTech, Spain, in 2016, a M.Sc. (Tech.) degree in ICT from the University of Turku, a M. Eng. degree in Electronics and Communication Engineering from Fudan University, China, in 2018, and a Ph.D. (Tech.) degree from the University of Turku in 2022. His research interests include multi-robot systems, collaborative



professor at the Hamburg University of Technology, Germany, from August 2015 to July 2017. He has more than 100 scientific publications on topics such as intelligent control, machine intelligence, nonlinear systems, robotics, and mechatronics.

Wallace Moreira Bessa is an Associate Professor of Mechanical Engineering at the University of Turku, Finland, and the Head of the Smart Systems Lab. Previously, he was an Associate Professor at the Federal University of Rio Grande do Norte, Brazil (2008-2021), and an Assistant Professor at the Federal Center for Technological Education, Brazil (2004-2008). In 2014, he was awarded the prestigious Humboldt Research Fellowship for Experienced Researchers, which allowed him to be a visiting



Salma Salimi is doctoral researcher at the Turku Intelligent Embedded and Robotic Systems (TIERS) Group at University of Turku, Finland since 2021. She received her B.Sc. Degree in Electrical Engineering-Electronics from University of Zanjan, Iran in 2016, and her M.Sc. Degree in Electrical Engineering-Control, from University of Tabriz, Iran in 2019. Her research interests include multi-robot systems, multi-robot collaboration and distributed ledger technologies



Sahar Salimpour is doctoral researcher at the Turku Intelligent Embedded and Robotic Systems (TIERS) Group at University of Turku, Finland. She received her MSc degree in Telecommunication engineering from the University of Tabriz, Tabriz, Iran in 2019. Her main research interests include robotics, deep learning, computer vision and machine learning.



operate and can perform a variety of functions in different fields like environmental monitoring, delivery, agriculture, and search and rescue missions. Prof. Westerlund is a senior member of IEEE.

Tomi Westerlund is a professor in robotics and autonomous systems. He is the research group leader of the Turku Intelligent Embedded and Robotics Systems (TIERS) lab at the University of Turku (UTU), Finland. TIERS's main research interests are in multi-robot systems, autonomous robots and edge computing. The main application areas for TIERS include collaborative and heterogeneous multi-robot systems, localization and mapping of urban and unstructured environments, hardware accelerators for AI and