

Improving Human-Robot Interaction Effectiveness in Human-Robot Collaborative Object Transportation using Force Prediction

J. E. Domínguez-Vidal and Alberto Sanfeliu

Abstract—In this work, we analyse the use of a prediction of the human’s force in a Human-Robot collaborative object transportation task at a middle distance. We check that this force prediction can improve multiple parameters associated with effective Human-Robot Interaction (HRI) such as perception of the robot’s contribution to the task, comfort or trust in the robot in a physical Human Robot Interaction (pHRI). We present a Deep Learning model that allows to predict the force that a human will exert in the next 1 s using as inputs the force previously exerted by the human, the robot’s velocity and environment information obtained from the robot’s LiDAR. Its success rate is up to 92.3% in testset and up to 89.1% in real experiments. We demonstrate that this force prediction, in addition to being able to be used directly to detect changes in the human’s intention, can be processed to obtain an estimate of the human’s desired trajectory. We have validated this approach with a user study involving 18 volunteers.

Index Terms—Physical Human-Robot Interaction, Object Transportation, Human-in-the-Loop, Force Prediction

I. INTRODUCTION

Robotics research has improved the behavior and abilities of robots extensively, enabling them to perform increasingly complex tasks with increasing accuracy. However, when the robot must work together with a human and a Human-Robot Interaction (HRI) occurs, the human is often a source of uncertainty that hinders the robot’s decision making.

In order to solve this problem, different methods have been developed to estimate the human’s preferences [1], [2] or their intention (which among several items they will choose [3], which action they will perform [4], which trajectory they will describe [5]–[7]...), usually focusing on specific tasks. Thus, in a collaborative search we want to know in which area the human is going to search [8], in a handover task we want to know the delivery point [9] or the path the human will follow [10] and in a collaborative transport task, where the human wants to take the object or how fast they wants to move [11].

Focusing on the latter task, in this work we take a different approach to that usually found in the literature. Instead of using a controller or some way of obtaining the speed with which the human wishes to move, we take advantage of the fact that this is a task involving physical contact and,

Work supported under the European project CANOPIES (H2020-ICT-2020-2-101016906) and the MINECO/AEI ROCOTRANSP project (PID2019-106702RB-C21 / AEI /10.13039/501100011033) and by JST Moonshot R & D Grant Number JPMJMS2011. The first author acknowledges Spanish FPU grant with ref. FPU19/06582.

The authors are with the Institut de Robòtica i Informàtica Industrial (CSIC-UPC). Llorens Artigas 4-6, 08028 Barcelona, Spain. {jdominguez, sanfeliu}@iri.upc.edu. The first one is the corresponding author.

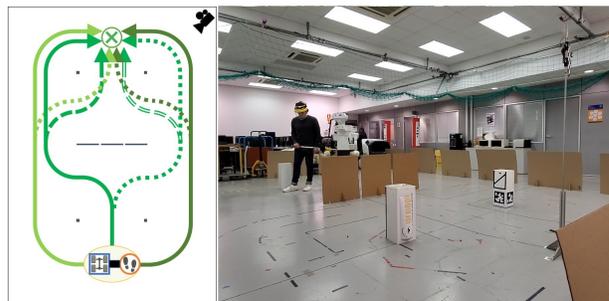


Fig. 1. **Human-robot pair transporting an object.** Both agents must navigate through a complex environment with multiple walls and columns to transport an aluminium bar until the goal *Left* - Designed setup. At least eight possible routes. Camera icon represents picture’s point of view on the right. *Right* - Real setup. Goal marked with a chequered flag.

therefore, the transfer of information between the two agents takes place through the exchange of forces. Thus, we develop a predictor of the force to be exerted by the human during the next second based on the force recently exerted and different information from the environment. We believe that this prediction can serve as a good source of information as it can be used to detect when the human’s intention disagrees with that of the robot, the speed with which they wants the task to be executed or the trajectory they want to follow.

Specifically, this article is a continuation of our previous work [12], in which we used a collaborative transportation task to present our Perception-Intention-Action (PIA) cycle although without exploring the advantages of having a prediction of the human’s future intention. In this work, we present our force predictor and use it in a first round of experiments (see Fig. 1) to condition the robot’s planner, so that its movements take into account not only the current force being exerted by the human but also the estimation of the force they will make.

In the remainder of the article, Section II presents the related work. Section III presents our force predictor, including its architecture based on known Deep Learning models and how the dataset used to train it was obtained. Section IV shows the results obtained both regarding the performance of the predictor and its effect on HRI through real experiments. Finally, Section V presents the conclusions.

II. RELATED WORK

Tasks originally involving manipulation and later human-robot collaborative transport have generally been analyzed with purely control-based approaches. Admittance control systems are common [13], both directly [14] or using two-

level control schemes [15], and obtaining the reference of their admittance controller from an equilibrium trajectory [16] or even visual servoing [17]. Their complementary implementations of the same control goal, i.e., impedance controllers, are also common [18]–[20]. However, in general, these solutions are based on adapting to the human in the best way and as fast as possible, but they do not seek to be anticipatory or proactive, as this would require a prediction of the human’s future behavior.

Other works do rely on some form of prediction to improve their performance. In [21] and [22], they use a prescribed performance estimation of the human’s desired trajectory for the transported object as an input of an impedance controller. A similar approach is followed in [23] where they make a Bayesian estimation of both human’s impedance and their motion intention as extra inputs to an adaptive impedance control scheme.

Focusing on models based on Deep Learning, in [24] and later in [25] they use a Radial Basis Function Neural Network to estimate the trajectory that the human’s limb wishes to follow and even to adapt to sudden changes in the desired trajectory. Meanwhile, [26] uses Reinforcement Learning to estimate and adapt in real-time the damping coefficient for a variable admittance controller. Learning from Demonstration (LfD) is used in [27] where Gaussian Mixture Models are the workhorse to learn a model of the task and then integrate it in an impedance controller. [11] also uses LfD to allow the robot to predict the speed profile that the human would like to follow to complete the task, in this case using Weighted Random Forrest.

While some work like [24] or [11] use the force exerted by the human as an input, to the best of our knowledge there is no work that predicts the force to be exerted by the human rather than the speed or the position to which the human wants to carry the object. Even less so if it is a task in which the transported object has to be moved over a middle distance while avoiding obstacles on the way and being able to change route on the fly. We consider that this force prediction can be a very valuable source of information since it is the dominant form of communication in this type of task and allows subsequently estimating the speed or position desired by the human. We use the force that the human has exerted recently as in [11], but we also seek for inspiration in [28]–[30] to create a model that combines convolutional and Long Short-Term Memory (LSTM) layers while considering more contextual information than either of them and using it to predict the next force to be exerted instead of the human’s movement.

III. HUMAN FORCE PREDICTOR FOR COLLABORATIVE OBJECT TRANSPORTATION

As mentioned above, in a collaborative human-robot transport task, both agents communicate mainly through the force they exchange. Therefore, being able to predict the force that the human is going to exert can serve as input to subsequently understand their intention and detect the changes that may occur during the task. To visualize this, we will use as a use

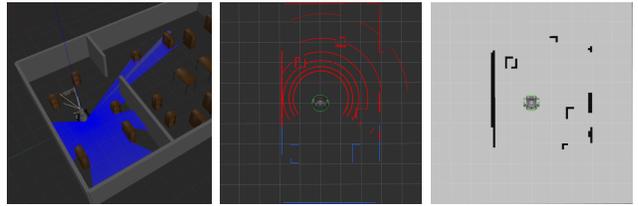


Fig. 2. **Occupancy map from LiDAR.** *Left* - Simulation in Gazebo of a robot in an indoors environment. *Middle* - Environment detection with front LiDAR (in red) and rear LiDAR (in blue). *Right* - Occupancy map with the obstacles detected. Pixel in black if there is an obstacle in the corresponding 10x10 cm equivalent cell.

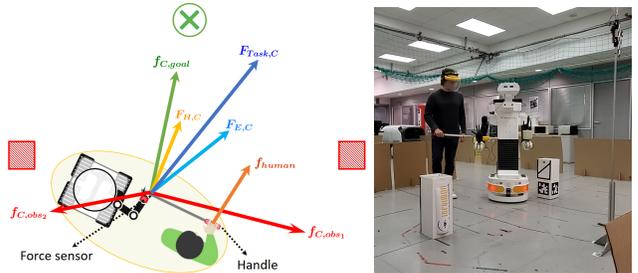


Fig. 3. **Forces involved in the considered collaborative transportation task.** *Left* - Main forces considered. Goal attractive force in green, obstacle repulsive force in red, environment force in light blue, human real force in orange, human’s measured force in yellow and task force in dark blue. *Right* - Example of real experiment at the same situation. Robot avoids corner and human impulses towards the goal. The object is an aluminium bar.

case a task where human and robot must transport a light object a middle distance of 6–8 m with walls and columns in the way and where there are multiple possible routes.

A. Problem definition

To obtain our prediction of the force to be exerted by the human, we process five sources of information.

First, we need information about the environment in which the task takes place. This environment is perceived in two ways. On the one hand, we use the LiDAR/LaserScan present in virtually all mobile robots to obtain an occupancy map around it. This map has a size of 100x100 pixels and each pixel represents whether or not the equivalent 10x10 cm cell in the real world is occupied by an obstacle (see Fig. 2). This image can be processed with a convolutional neural network (CNN) pipeline to extract the most representative features. This feature extraction is achieved through sequential convolution between the kernels at each layer and the feature maps produced in the preceding layer.

The second way to obtain information from the environment is to represent it by repulsive and attractive forces, taking advantage of the fact that the task is fundamentally governed by the exchange of forces between the two agents. Based on [31], we make each of the O obstacles detected by the robot to generate a repulsive force $f_{C,obs} \in \mathbb{R}^2$ that makes the robot move away from them. Likewise, a global planner generates a route to the task goal (known location to which the pair takes the object). This plan is split into a series of waypoints that generate an attractive force $f_{C,goal} \in \mathbb{R}^2$ so that the robot tends to follow this route. The combination

TABLE I
PERFORMANCE OBTAINED WITH DIFFERENT GRAPHIC CARDS

Graphic Card	Time Delay [ms]		FPS	
	(min./avg./max.)		(min./avg./max.)	
GTX 1060 Mobile (80 W)	70.1	- 71.8 - 75.2	13.3	- 13.9 - 14.3
GTX 1660 Ti Desktop	49.9	- 53.0 - 55.3	18.1	- 18.9 - 20.0
RTX 3060 Mobile (80 W)	41.2	- 43.8 - 46.1	21.7	- 22.8 - 24.3
RTX 3080 Ti Desktop	16.0	- 16.9 - 18.1	55.3	- 59.2 - 62.5

well as the speed commands delivered to the robot’s wheels. To increase the variability, a fifth scenario is designed with a higher number of possible routes (see Fig. 1) and 13 new volunteers (age: $\mu = 31.28$, $\sigma = 8.61$) perform the same collaborative transportation task, in this case twice each, and the same information is recorded as in the previous cases.

All the recorded sequences are split into sub-sequences of $N + T$ samples with an overlapping between sub-sequences of $(N+T)/2$ samples so that the model receives as input the first N samples and tries to predict the force exerted by the human in the last T samples. In total, 12239 sub-sequences are obtained, which are divided into the classic splits of training (90%: 11015 sub-sequences), validation (5%: 612 sub-sequences) and testing (5%: 612 sub-sequences) datasets.

The training is performed using Adam as the optimizer with its default parameters. Learning rate decay is used up to a minimum $lr = 3 \times 10^{-5}$ with a decay factor of 0.96. The model is trained for a maximum of 65 epochs, although using early stopping to avoid overfitting. Training was not observed to exceed epoch 60. An NVIDIA RTX 2080 Ti graphics card was used, training for 75-80 minutes.

IV. RESULTS

Once our force predictor has been trained, we first check its ability to correctly predict the force that the human will exert during the next second using the testset split. After that, we perform a new round of experiments in which 18 new volunteers (age: $\mu = 29.44$, $\sigma = 7.67$) perform the same collaborative transport task in the fifth scenario developed, as this is the scenario with the most possible routes and therefore the most demanding for the predictor. This new round of experiments allows us to test the generality of our predictor with different candidates who may have different preferences. In addition, it also allows us to perform a first user study of how the use of our predictor can improve different parameters associated with an effective HRI. All the experiments reported in this work have been performed under the approval of the ethics committee of the Universitat Politècnica de Catalunya (UPC)¹ in accordance with all the regulations and relevant guidelines (ID: 2021.10).

¹UPC ethics committee: <https://comite-etica.upc.edu/en>

TABLE II
EVOLUTION OF MEAN ERROR AND PERCENTAGE OF CORRECT PREDICTIONS IN TESTSET

Measure	Time [ms]			
	100	300	500	1000
Error F_x [N]	0,210	0,243	0,249	0,282
Error F_y [N]	0,101	0,126	0,126	0,151
Error $ F $ [N]	0,226	0,269	0,273	0,313
Error $\angle F$ [°]	6,731	7,081	7,132	7,527
Error $ F < 1.2$ N &	94,4	93,5	93,3	92,3
Error $\angle F < 18^\circ$ [%]				

TABLE III
EVOLUTION OF MEAN ERROR AND PERCENTAGE OF CORRECT PREDICTIONS IN REAL EXPERIMENTS

Measure	Time [ms]			
	100	300	500	1000
Error F_x [N]	0,282	0,312	0,316	0,352
Error F_y [N]	0,152	0,165	0,168	0,182
Error $ F $ [N]	0,314	0,344	0,348	0,395
Error $\angle F$ [°]	7,547	7,863	7,903	8,296
Error $ F < 1.2$ N &	92,2	91,2	90,8	89,1
Error $\angle F < 18^\circ$ [%]				

For this purpose, our predictor is encapsulated in a ROS (Robot Operating System) node that receives the necessary information flows, formats them and delivers them to the trained model so that it can deliver its prediction of the force to be exerted. This prediction is used to condition the robot’s global planner so that it takes the prediction into account when choosing the route to follow. Thus, each new volunteer performs the task twice, once without the predictor and once with the predictor conditioning the robot’s planner. After each execution, they fill out a short questionnaire to analyze their perception of the interaction².

For the sake of transparency, this ROS node also allows us to test the response speed of our predictor using different graphics cards. Table I shows a summary of the time it takes to execute the line of code that calls our model and generates a new prediction once a new sample of all information inputs has been received. Since we have tested with cards corresponding to different generations, we could not use the same drivers in all of them. Instead, the most recent driver as well as the latest version of CUDA compatible with each driver has been used with each one. As can be seen, even a 2016 laptop card can deliver the desired 10 FPS for this task. For the real experiments reported in this section, an MSI GS66 laptop with an RTX 3060 (80 W) was used.

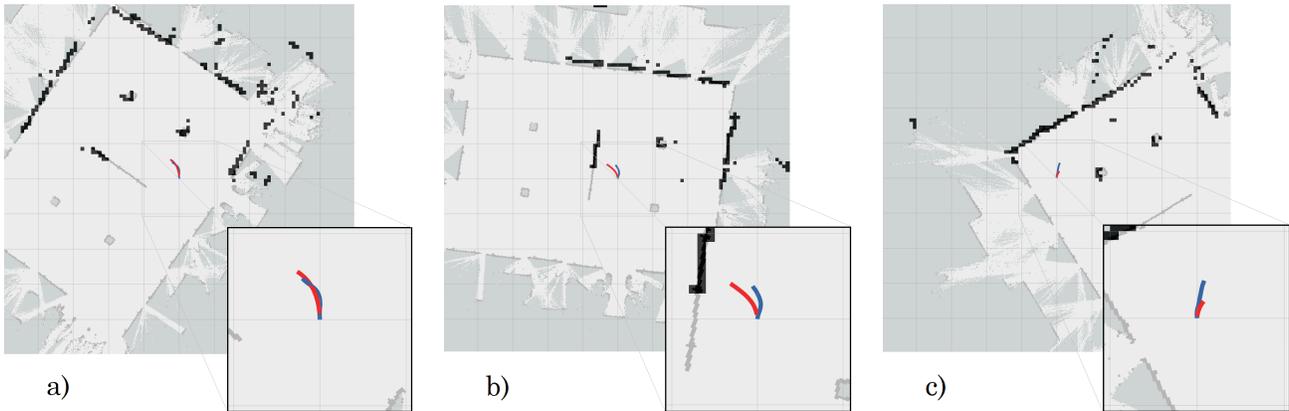


Fig. 5. **Trajectory estimation and real trajectory for 1 s in different situations.** Estimation in red and actual trajectory in blue. A - Correct estimation in normal situation. B - Bad estimation of trajectory (even with correct force prediction) when human’s force goes against normal operation, in this case, over-avoiding the obstacle in the left. C - Bad estimation of trajectory (even with correct force prediction) when human is exerting an extremely low force.

TABLE IV

COMPARISON OF MEAN ERROR ESTIMATING HUMAN TRAJECTORY WITH DIFFERENT MODELS. * MARKS VALUES OBTAINED BY INTERPOLATION FROM LAPLAZA EL AL. [10].

Model	L2 [m]	
	500 ms	1000 ms
Martinez el al. [6]	0.159*	0.317*
Mao et al. [7]	0.081*	0.161*
Laplaza et al.. [10]	0.072*	0.142*
2nd order polynomial	0.121	0.273
With our force predictor	0.092	0.198

A. Human Force Predictor Accuracy

We compute the absolute error in Cartesian coordinates between our predicted human force and the ground truth force for the same input sub-sequence. Table II contains the temporal evolution of the computed errors along the testset split. In addition to the error in each axis, we also calculate the absolute error between the modulus of the predicted and the actual force. The same calculation is performed with the angle. Finally, Table II also shows the percentage of samples that present an error in the modulus and angle of less than 10%, that is, $1.2 N$ in the modulus since the maximum force considered is $12 N$ and 18° since the angle moves between -180° and 180° .

As it can be seen, more than 90% of the samples remain within the desired threshold even after 1 s. It is observed that the main source of error is due to the angle. This is explained by the fact that abrupt changes in the force exerted by the human usually involve changes in the direction of the force rather than changes in its modulus.

The previously mentioned new round of experiments allows us to use the experiment in which the predictor is not used to obtain the same measurements but with candidates that did not appear in the training dataset. Table III shows

that the error made when estimating the force is higher in both axes. This results in a higher error in both the modulus and angle of the predicted force for all time intervals. This translates into a reduction of between 2.2% and 3.2% in the number of samples that fall within the previous threshold.

This reduction is due to two factors. First, the samples obtained in the previous experiments had been shuffled to generate the training and testing datasets so that, although the testing dataset was not seen by the model during training, its samples corresponded to the same volunteers. By using a different population sample, we can test the true generality of the model. Secondly, only the fifth scenario is being tested, which is more demanding for the predictor as more routes are available, whereas in the training and subsequent testing samples of all five scenarios were used, including the four used in our previous work [12] with fewer possible routes. Overall the model is shown to be quite general with a good performance in both testset and real experiments.

B. Force Predictor used for Movement Estimation

As discussed in section III-A, the combination in eq. (2) of the force exerted by the human, $F_{H,C}$, and the robot force due to its interpretation of the environment, $F_{E,C}$, is sent to a controller to generate the velocity commands that the robot follows. For simplicity, we will consider the force on the X-axis determines the linear velocity of the robot and the force on the Y-axis the angular velocity. Thus, the considered maximum force of $12 N$ on the X-axis would cause the maximum linear velocity of $0.65 m/s$ and this same force of $12 N$ on the Y-axis would cause the maximum angular velocity of $1 rad/s$. More details in [32].

Therefore, this prediction of the human’s force can be used to know the trajectory that the human would like to follow, i.e. to obtain an estimate of the intended motion. For this purpose, the force prediction can be given to the same controller and the resulting velocity can be integrated to obtain the estimated trajectory. It is worth mentioning that this estimation is only taking into account the force of the human, while the actual trajectory is determined by the

²Experiments example: <https://youtu.be/6kBZw1t3aEg>

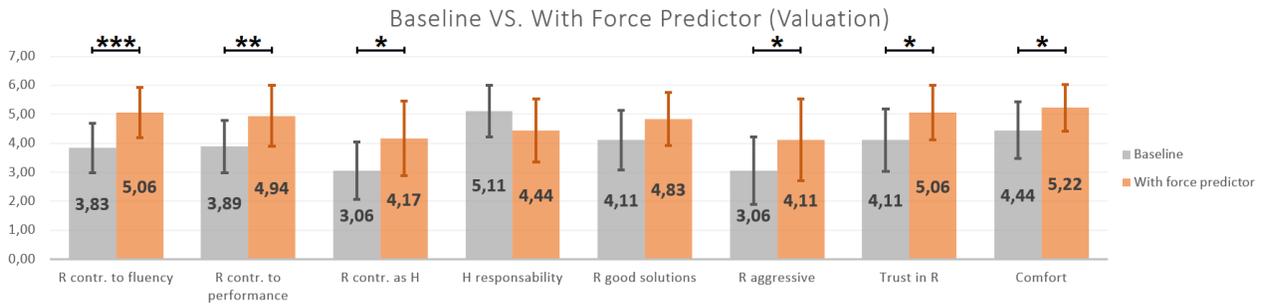


Fig. 6. **Assessment of different aspects involved in an effective HRI.** Valuation from 1 (very low) to 7 (very high) of the user experience. Statistical significance marked with *: $p < 0.05$, **: $p < 0.01$, ***: $p < 0.001$.

combination of the force exerted by both agents. This implies that this estimation will be correct as long as human and robot collaborate and interpret the environment in the same way. Fig. 5 - A shows an example of this case.

If the human chooses to follow a sub-optimal route or to avoid an obstacle at all costs even though this may lead the couple to collide with another, this prediction will not coincide with the trajectory that is finally followed (see Fig. 5 - B) but it will still constitute a valuable source of information to be able to interpret the human's intention demonstrating the usefulness and versatility of using a force predictor. A final case worth mentioning is when the human exerts a very slight force. The robot will complete the task by performing most of the effort so that prediction and actual trajectory will disagree (see Fig. 5 - C).

To get a qualitative idea of how good this motion estimation is, Table IV shows a comparison with other models of human motion estimation first and with a system that estimates the trajectory in this task by approximating a second order curve to the previously described trajectory. The values shown in the first three methods are obtained by interpolation of the values reported by Laplaza et al. in [10]. The comparison with these methods is not entirely fair since they are based on predicting the motion that the human will describe in tasks other than ours and typically having a camera pointed at the human instead of estimating it through the force exerted on an object as in our case. Anyway, they allow us to get an idea of what values the state of the art is moving at and show that our force predictor used to estimate the trajectory that the human-robot pair will follow does an acceptable job.

If we compare the results of this estimation made with our force predictor with one obtained by fitting a second degree curve to the trajectory described during the last second (the last 1 s is used instead of the last 2 s as in the case of our predictor so that this fit is more sensitive to the most recent variations), we can observe that even our course system is able to improve the result significantly.

C. Real Experiments User Study

This estimation of the human's desired trajectory can be used to condition the robot's global planner by making the following waypoints adapt to this estimation. This causes the attractive force in (1) to tend to go in the direction

desired by the human making the robot force in (2) more like the predicted force exerted by the human. This system is used in one of the two experiments performed by the 18 new volunteers. After each of the experiments, they fulfill a short questionnaire to assess different aspects of the HRI. Fig. 6 shows the results. All variables shown have passed a Saphiro-Wilk test for normality so that ANOVA tests can be applied to check for statistically significant variations (using the criterion of $p < 0.05$).

Looking at the robot's contribution to the task, we observe a statistically significant increase in both its contribution to interaction fluency (baseline $\mu=3.83$, $\sigma=0.86$; with force predictor $\mu=5.06$, $\sigma=0.87$; $p < 0.001$) and to task performance (baseline $\mu=3.89$, $\sigma=0.90$; with force predictor $\mu=4.94$, $\sigma=1.06$; $p=0.008$). Similarly, the subjective perception that the robot contributes to the task in equal proportion to the human also increases (baseline $\mu=3.06$, $\sigma=1.00$; with predictor $\mu=4.17$, $\sigma=1.30$; $p=0.014$). This results in a decrease in perceived human responsibility, that is, in how attentive the human must be for the task to be successfully executed, although not statistically significant (baseline $\mu=5.11$, $\sigma=0.90$; with predictor $\mu=4.44$, $\sigma=1.10$; $p=0.216$).

If we ask the human to rate the quality of the solutions proposed by the robot, the use of the predictor produces an increase although not statistically significant (baseline $\mu=4.11$, $\sigma=1.02$; with predictor $\mu=4.83$, $\sigma=0.92$; $p=0.215$). However, there is also a significant increase in perceived robot aggressiveness (baseline $\mu=3.06$, $\sigma=1.16$; with predictor $\mu=4.11$, $\sigma=1.41$; $p=0.038$). We believe this is because the robot reacts earlier to changes in the human's intention by detecting and taking them into account through the predictor.

Finally, a statistically significant increase is also observed in trust in the robot (baseline $\mu=4.11$, $\sigma=1.08$; with predictor $\mu=5.06$, $\sigma=0.94$; $p=0.025$) and in comfort running the task (baseline $\mu=4.44$, $\sigma=0.98$; with predictor $\mu=5.22$, $\sigma=0.81$; $p=0.035$) demonstrating the usefulness of our predictor to obtain an effective HRI even using it in a simple way as shown here.

V. CONCLUSIONS

We have approached the middle distance collaborative transport task from a different point of view by generating a force predictor instead of trying to predict the velocity profile or the human's desired trajectory. This predictor based on a

Deep Learning model has proven to provide predictions of up to 1 s with an acceptable error over 89% of the time in real experiments. This prediction of the force to be exerted by the human can be further processed and used to obtain an estimate of the human's intended trajectory. This opens the door to improve the estimates made in other works in the literature as well as to obtain a better understanding of the human's intention. This is due to the fact that this predictor allows to quickly detect changes of tendency in the force exerted by the human and, with it, their preferences regarding the route to follow, when to brake, when to accelerate or when to move more or less away from an obstacle.

Through a simple conditioning of the robot's planner to take this prediction into account, it has been proven through a user study that the use of our force predictor can improve multiple parameters associated with a successful HRI such as the perception of the robot's contribution to the task, comfort or trust in the robot.

ACKNOWLEDGMENTS

The authors want to express their most sincere gratitude to Ferrán Cortés and Patrick Grosch for their technical support and to all the volunteers for their infinite patience performing the experiments.

REFERENCES

- [1] E. A. Sisbot and R. Alami, "A human-aware manipulation planner," *IEEE Transactions on Robotics*, vol. 28, no. 5, pp. 1045–1057, 2012.
- [2] T. Kruse, A. K. Pandey, R. Alami, and A. Kirsch, "Human-aware robot navigation: A survey," *Robotics and Autonomous Systems*, vol. 61, no. 12, pp. 1726–1743, 2013.
- [3] C.-M. Huang and B. Mutlu, "Anticipatory robot control for efficient human-robot collaboration," in *2016 11th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, 2016, pp. 83–90.
- [4] H. S. Koppula and A. Saxena, "Anticipating human activities using object affordances for reactive robotic response," *IEEE transactions on pattern analysis and machine intelligence*, vol. 38, no. 1, pp. 14–29, 2015.
- [5] K. Fragkiadaki, S. Levine, P. Felsen, and J. Malik, "Recurrent network models for human dynamics," in *Proceedings of the IEEE international conference on computer vision*, 2015, pp. 4346–4354.
- [6] J. Martínez, M. J. Black, and J. Romero, "On human motion prediction using recurrent neural networks," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 2891–2900.
- [7] W. Mao, M. Liu, and M. Salzmann, "History repeats itself: Human motion prediction via motion attention," in *Computer Vision—ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part XIV 16*. Springer, 2020, pp. 474–489.
- [8] M. Dalmasso, A. Garrell, J. E. Domínguez-Vidal, P. Jiménez, and A. Sanfeliu, "Human-Robot Collaborative Multi-Agent Path Planning using Monte Carlo Tree Search and Social Reward Sources," in *IEEE Int. Conf. on Robotics and Automation (ICRA)*, 2021.
- [9] Z. Li and K. Hauser, "Predicting object transfer position and timing in human-robot handover tasks," *Science and Systems*, p. 38, 2015.
- [10] J. Laplaza, F. Moreno-Noguer, and A. Sanfeliu, "Context and Intention aware 3D Human Body Motion Prediction using an Attention Deep Learning model in Handover Tasks," *IEEE*, 2022, pp. 4743–4748.
- [11] A. Al-Yacoub, Y. Zhao, W. Eaton, Y. M. Goh, and N. Lohse, "Improving human robot collaboration through force/torque based learning for object manipulation," *Robotics and Computer-Integrated Manufacturing*, vol. 69, p. 102111, 2021.
- [12] J. E. Domínguez-Vidal, N. Rodríguez, and A. Sanfeliu, "Perception-Intention-Action Cycle as a Human Acceptable Way for Improving Human-Robot Collaborative Tasks," in *Proceedings of the 2023 ACM/IEEE International Conference on Human-Robot Interaction*, 2023, p. to appear.
- [13] S. Tarbouriech, B. Navarro, P. Fraise, A. Crosnier, A. Cherubini, and D. Sallé, "Admittance control for collaborative dual-arm manipulation," in *2019 19th International Conference on Advanced Robotics (ICAR)*. IEEE, 2019, pp. 198–204.
- [14] K. Kosuge and N. Kazamura, "Control of a robot handling an object in cooperation with a human," in *Proceedings 6th IEEE International Workshop on Robot and Human Communication. RO-MAN'97 SENDAI*. IEEE, 1997, pp. 142–147.
- [15] O. M. Al-Jarrah and Y. F. Zheng, "Arm-manipulator coordination for load sharing using reflexive motion control," in *Proceedings of International Conference on Robotics and Automation*, vol. 3. IEEE, 1997, pp. 2326–2331.
- [16] A. Bussy, P. Gergondet, A. Kheddar, F. Keith, and A. Crosnier, "Proactive behavior of a humanoid robot in a haptic transportation task with a human partner," in *2012 IEEE RO-MAN: The 21st IEEE International Symposium on Robot and Human Interactive Communication*. IEEE, 2012, pp. 962–967.
- [17] D. J. Agravante, A. Cherubini, A. Bussy, and A. Kheddar, "Human-humanoid joint haptic table carrying task with height stabilization using vision," in *2013 IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE, 2013, pp. 4609–4614.
- [18] D. J. Agravante, A. Cherubini, A. Bussy, P. Gergondet, and A. Kheddar, "Collaborative human-humanoid carrying using vision and haptic sensing," in *2014 IEEE international conference on robotics and automation (ICRA)*. IEEE, 2014, pp. 607–612.
- [19] Z. Li, J. Liu, Z. Huang, Y. Peng, H. Pu, and L. Ding, "Adaptive impedance control of human-robot cooperation using reinforcement learning," *IEEE Transactions on Industrial Electronics*, vol. 64, no. 10, pp. 8013–8022, 2017.
- [20] X. Yu, B. Li, W. He, Y. Feng, L. Cheng, and C. Silvestre, "Adaptive-constrained impedance control for human-robot co-transportation," *IEEE transactions on cybernetics*, vol. 52, no. 12, pp. 13 237–13 249, 2021.
- [21] C. N. Mavridis, K. Alevizos, C. P. Bechlioulis, and K. J. Kyriakopoulos, "Human-robot collaboration based on robust motion intention estimation with prescribed performance," in *2018 European Control Conference (ECC)*. IEEE, 2018, pp. 249–254.
- [22] K. I. Alevizos, C. P. Bechlioulis, and K. J. Kyriakopoulos, "Physical human-robot cooperation based on robust motion intention estimation," *Robotica*, vol. 38, no. 10, pp. 1842–1866, 2020.
- [23] X. Yu, W. He, Y. Li, C. Xue, J. Li, J. Zou, and C. Yang, "Bayesian estimation of human impedance and motion intention for human-robot collaboration," *IEEE transactions on cybernetics*, vol. 51, no. 4, pp. 1822–1834, 2019.
- [24] S. S. Ge, Y. Li, and H. He, "Neural-network-based human intention estimation for physical human-robot interaction," in *2011 8th International Conference on Ubiquitous Robots and Ambient Intelligence (URAI)*. IEEE, 2011, pp. 390–395.
- [25] Y. Li and S. S. Ge, "Human-robot collaboration based on motion intention estimation," *IEEE/ASME Transactions on Mechatronics*, vol. 19, no. 3, pp. 1007–1014, 2013.
- [26] F. Dimeas and N. Aspragathos, "Reinforcement learning of variable admittance control for human-robot co-manipulation," in *2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2015, pp. 1011–1016.
- [27] E. Gribovskaya, A. Kheddar, and A. Billard, "Motion learning and adaptive impedance for robot control during physical interaction with humans," in *2011 IEEE International Conference on Robotics and Automation*. IEEE, 2011, pp. 4326–4332.
- [28] Y. Cheng, L. Sun, C. Liu, and M. Tomizuka, "Towards efficient human-robot collaboration with robust plan recognition and trajectory prediction," *IEEE Robotics and Automation Letters*, vol. 5, no. 2, pp. 2602–2609, 2020.
- [29] K. Simonyan and A. Zisserman, "Two-stream convolutional networks for action recognition in videos," *Advances in neural information processing systems*, vol. 27, 2014.
- [30] Q. Xiong, J. Zhang, P. Wang, D. Liu, and R. X. Gao, "Transferable two-stream convolutional neural network for human action recognition," *Journal of Manufacturing Systems*, vol. 56, pp. 605–614, 2020.
- [31] D. Helbing and P. Molnar, "Social Force Model for Pedestrian Dynamics," *Physical review E*, vol. 51, no. 5, p. 4282, 1995.
- [32] J. E. Domínguez-Vidal, N. Rodríguez, R. Alquézar, and A. Sanfeliu, "Perception-Intention-Action Cycle in Human-Robot Collaborative Tasks," *arXiv preprint arXiv:2206.00304*, 2022.