Article

Data-Driven Modelling of Human-Human Co-Manipulation Using Force and Muscle Surface Electromyogram Activities

Ali Al-Yacoub *, Myles Flanagan, Achim Buerkle, Thomas Bamber, Pedro Ferreira, Ella-Mae Hubbard and Niels Lohse

Abstract: With collaborative robots and the recent developments in manufacturing technologies, physical interactions between humans and robots represent a vital role in performing collaborative tasks. Most previous studies have focused on robot motion planning and control during the execution of the task. However, further research is required for direct physical contact for human-robot or robot-robot interactions, such as co-manipulation. In co-manipulation, a human operator manipulates a shared load with a robot through a semi-structured environment. In such scenarios, a multi-contact point with the environment during the task execution results in a convoluted force/torque signature that is difficult to interpret. Therefore, in this paper, a muscle activity sensor in the form of an electromyograph (EMG) is employed to improve the mapping between force/torque and displacements in co-manipulation tasks. A suitable mapping was identified by comparing the root mean square error amongst data-driven models, mathematical models, and hybrid models. Thus, a robot was shown to effectively and naturally perform the required co-manipulation with a human. This paper’s proposed hypotheses were validated using an unseen test dataset and a simulated co-manipulation experiment, which showed that the EMG and data-driven model improved the mapping of the force/torque features into displacements.

Keywords: human-robot collaboration; human-human co-manipulation; data-driven modelling; mathematical modelling; object manipulation; impedance control

1. Introduction

Robots in the industry are starting to transfer from confined spaces into areas that are shared with humans, which reduces the operational cost for several industrial applications [1]. The co-existence of the human and the robot, however, raises many critical challenges regarding the safety of the human, tasks scheduling, and system evaluations [2]. To tackle these challenges, researchers in human-robot collaboration (HRC) focus on improving the production efficiency, safety, and collaboration quality between humans and robots.

Until now, the human often remains in a superior guidance role, which is due to human perception, cognition, and dexterity exceeding the capability of robots [3]. Robots, on the other hand, cope well with high payloads and repetitive tasks, while delivering high precision. Henceforth, several approaches established a combination of both human cognitive and perceptual abilities with the robot’s endurance to perform a collaborative task. This included intended physical contact between humans and robots during actions, such as hand-overs, co-manipulation, co-drilling, and many other applications [4,5].

To ensure human health and safety in such scenarios, robots operate at limited speeds and torque settings and perform a full stop in the case of a collision [6]. According to [5], however, these collisions are permissible when limiting the impact forces. Therefore,
another group of researchers proposed to control unavoidable collision by limiting the impact forces, such as the work presented by [7]. These methods rely on the kinematic and dynamic behaviour of the robot and human during the task. Such approaches require accurate models of the given setup [7–9]. Thus, there is a growing interest in measuring physical interactions between humans and robots [10]. Moreover, these interfaces have been utilised to establish intuitive human-robot interactions.

This paper presents a novel approach based on human-human co-manipulation to teach an industrial robot how to react to a human leader in a co-manipulation task. The main benefit of such an approach is the intuitive modelling that allows the robot to have similar behaviour to the human. This also allows the robot to isolate human haptic input. Hence, it is possible to estimate the force originating from the human guidance and the forces caused by contact with the environment (third contact point).

This paper’s proposed approach is based on two hypotheses investigated and validated on unseen test datasets. The first hypothesis includes adding a muscle activity sensor, namely an electromyograph (EMG), to improve the data-driven model quality. The second hypothesis is that data-driven approaches can achieve higher accuracy compared with mathematical modelling and hybrid modelling approaches, despite adding the complexity of human data processing. In order to validate the hypotheses, a simple leader–follower demonstration scenario was conducted. It provided the required data to model the follower behaviour and, thus, allowed teaching the robot to react similarly.

The paper is structured as follows; a brief literature review is presented in Section 2, Section 3 describes the co-manipulation problem from a mathematical point of view. Then, Section 4 details the equipment utilised during the human-human demonstration and data collection. The adopted research methodology in this research is outlined in Section 5, followed by the simulation setup and control-loop scheme in Section 6. Section 7 presents the results and discussion of the obtained outcome from the conducted experiment. Finally, the conclusions and future work are drawn in Section 8.

2. Literature Review

HRC is widely considered a critical concept to advance the quality and performance in several domains [9]. However, despite intensive research in this field, there are still many unsolved challenges that must be tackled to fully establish a safe and effective collaboration [11]. This includes vital questions such as: How can a robot predict human intentions? What role should the robot perform? At what time? Finally, how can a HRC setup be evaluated [12]? Whereas some companies have invested in HRC and have collaborative robots (cobots) on their manufacturing lines; many others are still waiting for more mature solutions.

The authors in [13] claimed that the reason behind this is the lack of knowledge when integrating cobots into manufacturing and business. Therefore, ref. [13] examined current training programs that inform manufacturers about finances and tools that support decision makers to analyse assembly workstations and determine whether HRC would be beneficial to their applications. The study concluded with the fact that there is a lack of knowledge about the integration of cobots into the manufacturing processes. This could also be due to several definitions of HRC being available. In some cases, HRC is treated as a synonym for human-robot interaction in general.

In order to distinguish different kinds of human-robot interactions, ref. [14] established criteria, such as a workplace, working time, aim, and contact. At the lowest level, human-robot coexistence includes a shared working environment, where tasks of the human and the robot do not interfere [15]. In addition to a shared working environment and working time, human-robot cooperation also includes a shared aim of the overall task [14]. On the highest level in HRC, physical contact between humans and robots is also permitted, which makes it the most challenging method of interaction among the three [15]. Hence, human-robot mutual awareness is required, as well as the exact timing of tasks and activities [11].
To understand the required mutual awareness, ref. [16] introduced a generic definition of several types of human-human and human-robot interactions, in which interaction was considered as a function of physical distances between the human and the robot. Therefore, these interactions were categorised into avoiding, passing, following, approaching, and touching. Hence, the parties in a co-manipulation context can be two humans, a human and a robot, or two robots who manipulate a shared object from point \( A \) in the workspace to point \( B \). According to [16], this concept can be extended to multiple humans with a single robot (multiple–single) or multiple humans with multiple robots (multiple–multiple).

Using this analogy, therefore, human-human co-manipulation can provide useful insights on how humans perform a task and consequently teach robots to perform such tasks with humans. Subsequently, various controllers have been designed to allow multiple robots to work together when manipulating an object [17]. Other researchers proposed controllers that are inspired by human anatomy and behaviour, and such research takes the damping and impedance characteristics of human movement into consideration [18]. The authors in [19] stated that human-friendly robot cooperation requires an adaptive impedance control that adjusts the robot impedance based on human characteristics. Hence, motion data from a human-human co-manipulation experiment was conducted.

Using dynamic model identification techniques, damping and stiffness factors have been estimated based on a simplified mathematical model (MM). The method mentioned above, however, requires an accurate MM to be conceived, which is often complicated and time-consuming. The collected data only contains position, speed, and acceleration, while the forces are estimated accordingly. Another approach to model human-human co-manipulation, similar to [18,19], can be found in [20]. An analysis was performed on the leader and follower human, in which one human provided the haptic guidance and the other human followed the haptic clues. The study concluded that most forces originated from the leader, while the follower focused on tracking the co-manipulated object.

A control scheme that enables a humanoid robot to perform a co-manipulation task with a human partner was proposed by [21]. The proposed control scheme consisted of three phases. At first, the robot predicted the human’s intentions to move based on a Force/Torque (F/T) sensor attached to the robot. In a second phase, the human and the robot switched roles, where the robot provided guidance and the human followed. In the third phase, the robot was controlled remotely with a joystick. The main problem with this work that it ha a discrete nature, and it required initiation through a joystick.

Since HRC requires the human and the robot to be co-existing in the same workspace, the design and development of a suitable HRC space is still an open challenge [22]. The authors in [23] examined the HRC shared workspace requirements based on a case study of disassembling press fitted components using collaborative robots. The study concluded that the compliance behaviour of a collaborative robot enabled the operator to work closely with the robot, which means lower installation costs and increasing efficiency as the robot and the human can work simultaneously.

However, many challenges need to be addressed regarding the performance evaluation, task assignments, and role management. Researchers proposed to equip the workspace with sensors that can improve the communication between the human and the robot to address the workspace challenges. These sensors can be classified as contextual sensors, such as cameras, motion trackers, biomechanical and psychological sensors, and motion sensors. These combined sensors are believed to improve human-robot communication as they can be used to infer the physical and psychological states of the human during the execution of the HRC task.

In more recent years, approaches have utilised Machine Learning (ML), including advanced classifiers such as Artificial Neural Networks (ANNs), to process sensory data [24]. In [25], an ANN was employed in human-human co-manipulation to predict the required motion based on the F/T input from the leader. The controllers used in this approach followed the desired trajectory to some degree of accuracy. However, the predicted trajectory was described as jerky, and the maximum error reached was 0.1–0.2 m. In further
research, ref. [26] investigated human-human co-manipulation and proposed the use of a Deep Neural Network (DNN) to accurately estimate the velocity and acceleration of the human over 50 time steps of 0.25 s. The trained model showed higher accuracy in comparison with the work presented in [25], although it was prone to noise interference.

In addition to haptic interfaces and F/T sensors, a growing interest in wearable sensors can be observed, which is intended to improve communication between humans and robots [27]. These approaches also aim to improve the mutual awareness of both parties in HRC [28]. Combining wearable sensors with existing technologies is believed to significantly improve HRC, as they can be used to infer the physical and psychological states of the human during the execution task [29,30].

However, processing signals originating from wearable sensors is considered challenging due to large amounts of subject-specific noise within the data. For example, EMG signal depends on the individual internal structure, such as skin temperature, blood flow rate, muscle structure, and many other factors [31]. Moreover, signals can even vary during different recording sessions for the same individual [32]. Nevertheless, advanced ML classifiers can be deployed to demonstrate their characteristic strengths of coping well with noise [33].

A common wearable sensor for detecting muscular fatigue, as well as applied muscular contraction, can be found in EMGs. The EMG signal directly correlates with the forces applied [34]. For instance, ref. [35] proposed the use of EMG to estimate the required stiffness during HRC tasks since stiffness is an essential component in cooperative leader-follower tasks. The estimated stiffness was employed in a hybrid force/impedance controller. EMG signals in conjunction with an Online Random Forest were used to detect muscular fatigue/or a human operator struggling with a high payload in order to trigger an assistance request for a cobot [27].

Additionally, Brain-Computer Interfaces (BCIs) have gained research interest in HRC. In one approach, a BCI was utilised for communicating human movement intentions to a robot [36]. A DNN was deployed to process electroencephalogram (EEG) signals, which allowed for measuring and classifying human upper-limb movements up to 0.5 s prior to the actual movements. A similar result have also been reported in [37]. Another approach based on a DNN to predict human intentions to move a limb was presented by [38]. In this paper, the human limb position was estimated based on torque readings collected from a F/T sensor attached to the robot end-effector. In general, estimating human-intention-based different sensors is still at an early stage, and further research is required.

Overall, there is an eminent research interest in HRC, of which co-manipulation can be viewed as a small building block. Strategies for classical co-manipulation control vary between force-based and motion-based; however, almost all are limited as they require accurate mathematical modelling pre-knowledge about the desired trajectory. Nevertheless, based on the literature, promising potential towards more accurate models can be found in ML, including ANNs.

Advanced ML algorithms enable the classification of bio-sensory data, such as EMGs. Utilising EMGs in conjunction with F/T data to predict human motions for co-manipulation tasks is still at an early stage. This would allow the collaborative robot to learn a behaviour, based on human-human co-manipulation within different data-driven models. Moreover, the performance of such data-driven models can be compared with previous mathematical models and hybrid models. This paper provides a comprehensive study on human-human co-manipulation, including a data-driven model, which also considers EMG signals.

3. Problem Definition

The human-human co-manipulation problem can be defined as transporting an object by at least two humans through unstructured/structured environments. In such a task, the human (leader) uses his/her perception to navigate the surroundings while communicating with others (followers) using haptic forces and verbal and gesture clues. The main focus of this paper is the haptic communication between the leader and the follower. Figure
1 depicts the fundamental concept of the co-manipulation task, in which two human operators are carrying out a co-manipulation task of handling a shared weight. The influential factors in such a scenario are the muscle stiffness, resultant forces, object mass, and object displacements in the 3D space.

Figure 1. Human-human co-manipulation problem.

The problem at hand can be described as a one-to-one mapping that aims to map follower displacement, and directions with the muscle EMG signal in response to the leader F/T input. Formally, the input data are the F/T data and follower EMG signal, while the output is the displacement of the load in the 3D space. Equations (1) and (2) depict the mathematical definition of the mapping problem.

\[
\text{Input} = \begin{cases} 
F, & F \in \mathbb{R}^{6 \times m} \\
\text{emg}, & \text{emg} \in \mathbb{R}^{2 \times m}
\end{cases} 
\]

\[
\text{Output} = \begin{cases} 
xd, & xd \in \mathbb{R}^{1 \times m} \\
yd, & yd \in \mathbb{R}^{1 \times m} \\
zd, & zd \in \mathbb{R}^{1 \times m}
\end{cases} 
\]

Henceforth, the problem can be mathematically defined as finding the mapping between the input (Equation (1)) and the output (Equation (2)) as shown in Equation (3).

\[
M(F, \text{emg}) = \text{Input} \rightarrow \text{Output} 
\]

Such a problem can be considered as a regression problem [39] that can be solved using ML approaches to identify suitable mapping while minimising the error.

4. Experimental Setup and Data Collection

To test the proposed hypotheses in this paper, an instrumented load was used to collect data during the co-manipulation task, as shown in Figure 2. Two sEMG sensors were utilised, which were fitted on the arm and the forearm as illustrated in Figure 3. For the biceps muscle, the electrodes of the sensors must be aligned with the muscle axial, which was identified as shown in Figure 3a, while the reference electrode must be shifted away from the muscle axial.

Similarly, the sensor electrodes on the forearm muscle must be fitted on the forearm muscle, specifically on the brachioradialis muscle [40], while the reference electrode placed on the outside of the forearm close the bone side; as illustrated in Figure 3b. The Myoware sensor used in this paper has on-board functionality that permitted the reading of the rectified signal, making the signal suitable to be integrated with the presented setup. The signals were sampled using a 12-bit Analogue-To-Digital Converter with a \(\sim 85\) Hz sampling frequency. The instrumented load is a wooden board with a 5.0 kg weight attached to it in the centre. The following sensors were utilised:

1. A six-axis F/T (F/T) sensor [41].
2. Surface electromyography (sEMG), [42], which is worn by the follower human
3. Motion tracker markers: eight cameras—VICON Vantage 5, https://www.vicon.com/hardware/cameras/vantage/. 
The F/T, VICON, and sEMG sensors were sampled at different frequencies; therefore, these sensory data were synchronised using an adaptive algorithm implemented in the `messageFilter-ApproximateTime` tool (http://wiki.ros.org/message_filters/ApproximateTime). The bottleneck of this tool is the slowest signal, the sEMG. Thus, the synchronised data will almost have the same frequency as the sEMG, which is around (∼85 Hz).

![Experimental setup](image)

**Figure 2.** Experimental setup.

![Fitting sEMG sensors](image)

**Figure 3.** Fitting sEMG (Surface electromyography) sensors on (a) biceps and (b) forearm.

For the experiment presented in this paper, two participants were asked to co-manipulate the load described above while avoiding an obstacle within the workspace. Figure 4 depicts the floor plan for the experiment and the path that the leader and the follower had to follow. The participants were assigned a leader or follower role and were not allowed to communicate during the experiment. This prevents the participants from verbally sharing their intention. The follower was equipped with muscle activity sensors, which were integrated with the Robot Operating System (ROS) network [43].

The 5.0 Kg load, in this scenario, provided some resistance while manipulating the object to emulate a realistic scenario. Additionally, each participant was only permitted to use one arm whilst carrying the object. This enabled the F/T readings to be mapped to...
a single local reference point. The leader guided the manipulation within the workspace while avoiding the obstacle and towards the endpoint. At the same time, the follower reacted to the leader’s movements and mimicked his/her motions until the endpoint was reached. The experiment was designed in this way to replicate human-robot co-manipulation.

Figure 4. Floor plan, human-human.

Sensor Placement

The F/T sensor was placed on the follower side in the centre of the object. The correct EMG sensor placement is essential as the quality of the signal can be affected. As such, the sensor must be placed over the centre of the muscle as shown in Figure 3 [42]. Finally, the motion capture reference point was placed on the leader’s side in the object’s centre, while the F/T sensor was located between the leader’s handle and the load. Hence, a transformation between the F/T coordinate into the workspace coordinate is required. Figure 5 illustrates the required transformation; also, it can be noticed that the coordinate system of the F/T sensor in which the Z-axis is aligned with the handle axial and XY plane was parallel to the surface of the sensor.

Figure 5. Coordinate system transformations: F/T sensors into VICON system.

5. Methodology

Four different sets of features were used to predict the required displacements based on leader guidance to validate the proposed hypotheses. The fitted models were evaluated based on unseen test datasets using the root mean square error (RMSE). An overall displacement error (overall RMSE) was calculated to determine the resultant error \( \sqrt{\text{RMSE}_x^2 + \text{RMSE}_y^2 + \text{RMSE}_z^2} \). The employed feature sets are shown in Table 1.
Table 1. Feature sets utilised to fit data-driven models.

| Feature Set   | Features                                                                 | Set Dimension |
|---------------|--------------------------------------------------------------------------|---------------|
| Features 1 (F1) | normalised F/T, normalised EMG (arm/forearm), previous 3D displacements | R^{11}        |
| Features 2 (F2) | F/T, previous 3D displacements                                           | R^{9}         |
| Features 3 (F3) | F/T, EMG (arm/forearm), previous 3D displacements                         | R^{11}        |
| Features 4 (F4) | normalised F/T, EMG (arm/forearm), previous 3D displacements             | R^{9}         |

The F/T features were the F/T data \( \mathbf{F} = \{ F_x, F_y, F_z, T_x, T_y, T_z \} \in \mathbb{R}^6 \), EMG arm and forearm \( \mathbf{EMG} = \{ emg_{\text{arm}}, emg_{\text{forearm}} \} \in \mathbb{R}^2 \), respectively and previous 3D displacement, which is the displacement from the previous timestamp \( \Delta = \{ \delta x, \delta y, \delta z \} \in \mathbb{R}^3 \). Then, the performance of the fitted models on the unseen test dataset was calculated using RMSE. This allows for testing the impact of including EMG sensory data in such a context. Finally, the performance of the best data-driven model was compared against the performance of the MM and the hybrid model. Another important feature in the co-manipulation tasks is the time, in which the human intention to move the shared object at time point \( t \) not only depends on the F/T and EMG at the same point but also depends on the displacements at the previous timestamp \( t-1 \), as summarised in Equation (4).

\[
\Delta(t) = \mathbf{M}(\mathbf{F}(t), \mathbf{EMG}(t), \Delta(t-1)) \tag{4}
\]

where \( \mathbf{M}(-) \) is the mapping function of the given features to the intended displacement \( \Delta \). The displacement \( \Delta(t-1) \) is measured using the VICON system by comparing the Cartesian position at \( t \) with the Cartesian position at \( t-1 \).

5.1. Mathematical Modelling

The problem described in Section 3, can be simplified as a mass–spring–damper system, as shown in Figure 6. The human arms can be simplified as a spring–damper on both sides of the transported object in 3D space.

![Figure 6. Simplified mathematical system, where \( F_f, F_l \) are the follower and leader forces, respectively; \( K_f, K_l \) are the follower and leader muscle stiffness, respectively; \( b_f, b_l \) are the follower and leader damping factor, respectively; and \( M \) is the mass of the shared load.](image)

Using Newton’s second law \( \sum F = ma \), the problem in Figure 6 can be described as shown in Equation (5).

\[
m \Delta a(t) + (b_f + b_l) \Delta v(t) + (k_f + k_l) \Delta X = 0 \tag{5}
\]

The superposition concept can be used for further simplification, as the 3D space system can be split into three separated equations (vector form), as shown in Equation (6). Based on the experimental setup and the problem at hand, the resultant force is measured using the F/T sensor. Furthermore, the total mass of the moving object is known; thus, the model described in Equation (5) can be simplified by omitting the muscle stiffness and damping effect as shown in Equations (5) and (6). To determine the displacements based on the measured forces, Equation (6) can be rewritten as in Equation (7).
\[
\begin{bmatrix}
  f_x \\
  f_y \\
  f_z 
\end{bmatrix}
= \begin{bmatrix}
  a_{xs} \\
  a_{ys} \\
  a_{zs}
\end{bmatrix} \begin{bmatrix}
  m
\end{bmatrix}
\] (6)

\[
\begin{bmatrix}
  \Delta x_s \\
  \Delta y_s \\
  \Delta z_s
\end{bmatrix}
= \iint \begin{bmatrix}
  \frac{f_x}{m} \\
  \frac{f_y}{m} \\
  \frac{f_z}{m}
\end{bmatrix} \partial^2 t
\]
(7)

In Equation (7), \(\Delta x_s, \Delta y_s, \text{ and } \Delta z_s\) are the displacements in the F/T sensors. Therefore, they must be transformed into the VICON coordinate system using a transformation matrix that was calculated based on the alignment of the F/T sensor with respect to the VICON markers (Figure 2). The final MM is shown in Equation (8), where \(T^v_s\) is the transformation matrix from the F/T sensor into the VICON coordinate system.

\[
\begin{bmatrix}
  \Delta x_v \\
  \Delta y_v \\
  \Delta z_v
\end{bmatrix}
= T^v_s \begin{bmatrix}
  x_v \\
  y_v \\
  z_v
\end{bmatrix} = T^v_s \iint \begin{bmatrix}
  \frac{f_x}{m} \\
  \frac{f_y}{m} \\
  \frac{f_z}{m}
\end{bmatrix} \partial^2 t
\]
(8)

5.2. Model-Free Approaches: Data-Driven Models

Modern manufacturing systems are complicated since they integrate a wide variety of equipment that extends from machinery and automation equipment on the shopfloor up to cloud systems and data-acquisition tools. In addition, the complexity on the shopfloor level due to the fast development of communication is exponentially increasing, which means higher data flows between different elements on the shopfloor. Consequently, equipment has higher nonlinearities, disturbances, and uncertainties. Therefore, it is almost impossible to describe these complicated systems using conventional mathematical models, such as differential equations or statistical models, in real applications. Nonetheless, with the fast advancement of sensing technology and data collection technologies, data-driven modelling (DDM) becomes more feasible in comparison with mathematical modelling [44].

Based on the regression problem explained in Section 3, we propose the use of Linear Regression (LR), Random Forest (RF) regression, Boosted-Trees (BT), and Recurrent Neural Networks (RNN) as these methods represent state-of-the-art approaches [45]. LR is well-known for simplicity and its ease of use. In contrast, BT and RF are ensemble ML approaches that are powerful and have shown high performance on several datasets. Finally, an RNN as part of a Deep Learning Neural Network is utilised. Hence, in this paper, we compared the performance of each ML algorithm and the performance of the data-driven approaches with the mathematical and hybrid models. In the following subsection, the data-driven approaches are explained in the co-manipulation context.

5.3. Hybrid Modelling Approach (HM)

As described in the mathematical modelling section, mathematical models are often derived from the fundamental laws of physics, such as Newton’s laws of motion. Physical models extrapolate well by design and, therefore, are preferred for model-based control approaches. In realistic scenarios, however, modelling errors exist due to omitted dynamics, modelling approximations, lack of suitable friction models, backlash, or unmodelled elasticity in the mechanism. Classically, these problems can be tackled with assumptions, linearisation around an operation point, and enhancing the model parameters based on theoretical or experimental methods. In the case of a very complex mechanism, however, these solutions might not be feasible.

On the contrary, data-driven modelling approaches utilise the behavioural response of the system for different inputs and then extract a set of generic rules that describe the correlations amongst the inputs and outputs without omitting or simplifying the system. The main drawback of such systems is that they are not always interpretable as in
physical/mathematical modelling. The quality of the data-driven model depends on the size and quality of the collected data.

Furthermore, data can barely ever deplete all possible configurations. Thus, a hybrid model can be used that combines simplified MM with a data-driven error model [46]. The target is to capture the main physical attributes of the given system (e.g., the robot) while substituting for model approximations and inaccuracies. Hence, the hybrid model has a grey-box character due to the mixture of a physical and data-driven (black-box) error model. In this paper, we decided to model the error within a mathematical model, Section 5.1, using the best data-driven approach from the previous section. Then, the problem description can be now rewritten, as shown in Equation (9).

\[
\begin{bmatrix}
\Delta x_s \\
\Delta y_s \\
\Delta z_s 
\end{bmatrix}
= \int \int \begin{bmatrix}
\frac{L_x}{m} \\
\frac{L_y}{m} \\
\frac{L_z}{m}
\end{bmatrix} \partial^2 t + E(\text{emg}, F) 
\]

(9)

where \(E(\text{emg}, F)\) is the error between the real displacement measured using VICON and the estimated displacement using the mathematical models. The error function \(E(\cdot)\) can be seen as a regression problem that can be tackled using the best data-driven approach. Consequently, the final model is a combination of MM and data-driven models. For evaluation purposes, the hybrid model was compared with all the approaches above on unseen test data.

6. Simulation Setup

A simulated UR10 robot with a 5.0 Kg load was exploited to evaluate the different methods mentioned above. Hence, the output from these methods was used with a Proportional–Integral–Derivative (PID) controller, as depicted in Figure 7. In this control scheme, the output \(\text{disp}(t)\) can be seen as a feedforward control scheme, in which the prediction of error \(\text{disp}(t)\) can be added to the error from the previous action (feedback-loop).

The inner loop is also a position control loop that attempts to maintain a precise position given the prediction from the data-driven models combined with the position error. The outer loop can be seen as a force-based control loop in which the robot must react to human EMG and forces in a spring–damper manner. Therefore, this control scheme is an Impedance Controller.

The simulation setup composed of a Workstation that runs Ubuntu 18.04 (developed by Canonical Ltd.), ROS-Indigo (Developed by Willow Garage, Menlo Park, CA, USA) with 100 Hz simulation frequency and the models developed earlier. The PID parameters were experimentally chosen, and similar settings were implemented to test the methods highlighted earlier.

7. Results and Discussion

7.1. Results

During the human-human demonstration, two participants performed the co-manipulation task, while data, from the F/T sensor, EMG sensor, and VICON, were
The collected, filtered, and synchronised data points totalled 5125. The EMG signal was collected only from the follower; one could argue that the sEMG data are insufficient to draw a generalised solution. However, studies revealed that the pattern of sEMG was comparable amongst different individuals; and hence, magnitude normalisation allows us to generalise the findings of this paper [47].

Out of the 5125 data points, 4212 data points were used for training and validation and 513 data points were used for testing. The synchronised data frequency was ~80 Hz, in which sensory data were synchronised using an approximation time tool developed for ROS. The F/T sensory data were filtered using a low-pass filter with a cut-off frequency at 50 Hz. The VICON data were filtered using moving-window-average and the manual removal of anomaly data that occurred due to flip [48]. The collected data included the F/T signal, the Cartesian position of the co-manipulated object, and the sEMG muscle activity signal of the follower’s right arm. As highlighted in the problem definition section, the goal is to find the mapping between relevant features and the displacements on the Cartesian space.

For training the models, four different sets of features were used as illustrated in Table 1, and these sets can be summarised as follows: normalised F/T signals and normalised EMG signals (F1), F/T signals (F2), F/T signals and EMG signals (F3), and normalised F/T signals (F4). The main reason behind this choice of features was to test the proposed hypotheses that the use of EMG signals can improve the accuracy of the data-driven models.

Figure 8 shows the accuracy of each predicted displacement in the X, Y, and Z directions for the data-driven models. The best model shown in Figure 8d was the RNN model based on feature set F1, which had the lowest RMSE with about 0.025 m on all axes and about 0.045 m overall error. The remaining data-driven approaches did not show the same impact of including the EMG data as illustrated in Figure 8a–c.
In terms of accuracy, however, the RF models showed tangent accuracy to the RNN models with an overall RMSE of around 0.05 m. Finally, the BT and LR models had similar performance with an overall RMSE of around 0.07 m as depicted in Figure 8a–c. Another significant result is that sEMG features did not necessarily improve the quality of the data-driven models, especially in the LR and BT models, in which the performance was almost constant regardless of the feature set.

The general trend regarding the accuracy in the $X - Y - Z$ direction is that the RMSE in the $Y$ direction was higher than the RMSE in $X$ and $Z$ across all models and feature sets, as illustrated in Table 2. The $Z$ axis models had the lowest RMSE but were still very close to the $X$ axis RMSE values. This performance variation on the $X - Y - Z$ models is believed to be due to the quality of the VICON data, which could be degraded due to reflective objects, obstacles, and light variations.

Table 2. Accuracy (m) in $X - Y - Z$ directions.

| Model | $X$    | $Y$    |
|-------|--------|--------|
| LR    | 0.032  | 0.060  |
| RF    | 0.030  | 0.060  |
| BT    | 0.030  | 0.054  |
| RNN   | 0.026  | 0.037  |

Models with the lowest RMSE were chosen and compared with the MM and HM, as shown in Figure 9. This figure shows that the RMSE values for the MM were (0.75, 1.34, 1.0)m. The HM (Figure 10) RMSE values were (0.051, 0.056, 0.051)m, which is comparable to the data-driven approach. We propose that these results occur due to the unknown dynamics naturally originating from the human body that allow for an adaptive non-linear change of stiffness and compliance, which is not captured in the simplified MM.

Figure 9. The best RMSE scores for data-driven models vs. the mathematical models.

As the RMSE scores for the MM were very large, it is not easy to compare amongst different models in Figure 9. Therefore, Figure 10 illustrates the accuracy of the data-driven models in comparison with HM, where it is clear that the RNN models had the lowest error (overall error 0.025 m) while the rest of the data-driven models had an error range between (~0.065 m for BT models) and (~0.075 m for LR models). The RNN models had the lowest variations on all axes variations in comparison with other methods.

This result is of tremendous importance, as in human-robot co-manipulation, movement variation (fluctuation) might result in jerky movements that represent a safety issue, especially if the human is physically interacting with the robot. By a closer look at how
the models behave in $X - Y - Z$ directions, all models had relatively larger error along the $Y$ directions. This indicates the lack of variation along the $Y$ axis in the captured data. Another possible explanation is that the participants blocked the VICON cameras’ field, which reduced the quality of the data in a certain direction.

![Figure 10. Data-driven models and hybrid-modelling (HM).](image)

### 7.1.1. Simulation Results

The first 50 sampled points from the fourth trial were used to control a simulated setup as described in Section 6. The overall displacements of these 50 points were $\sim 3.5$ m, and the predicted displacements were used as a set point for a PID controller as illustrated in Figure 7. The displacement error per-sampling point of the best data-driven models in comparison with MM and HM is depicted in Figure 11. It is obvious from this figure that the data-driven models had an error less than ($\sim 6\%$), the HM model had an error of ($\sim 7.8\%$), and the MM had an error of ($\sim 11.9\%$).

This shows that data-driven models had higher accuracy in comparison with HM and MM. The error of data-driven models, however, appeared to be tangibly similar (on average $\sim 4\%$), in which LR and BT had the lowest error followed by RF and RNN, respectively. Since the co-manipulation and force control, in general, is a non-linear problem, these results come as a surprise given the results in the previous section. The explanation for these results could be that the non-linear behaviour was achieved through the control scheme (a feedforward–feedback loop).

Another essential aspect in the human-robot co-manipulation task is the interaction forces during the execution [49]. Hence, during the simulation, interaction forces based on a dynamical model of the load (5.0 Kg) were estimated, and then physical metrics, such as the work and kinetic energy, were calculated as shown in Figure 12. The results show that the force, work, and kinetic energy of MM and HM were very high (relatively) compared to the data-driven approach.

This not only means that it is challenging to do co-manipulation based on these approaches but also that it is not safe to do such a task due to the speed variation ($acc \neq 0$). Jerky movements with some impact force can cause injuries to the human during co-manipulation tasks. On the other hand, the data-driven approaches appeared to have much smoother movements as illustrated by the lower interaction force, work, and kinetic energy, which indicates that movements occurred at a constant speed ($acc = 0$) with less jerky movements.
7.2. Discussion

The results revealed that the RNN models with feature set F1 were the most accurate, followed by the RF, BT, and LR models, respectively. However, the RNN and the RF models had tangible results. This is particularly important when a data-driven approach is adopted for a HRC system with limited computational resources. In that case, the RF models can give very accurate results with much lower computational requirements. Another important observation is that the RNN and RF models trained on normalised F/T and EMG (F1) had the best performance amongst other models. However, RF does not require normalised features, and the performance of RF models must be the same with and without normalisation. We speculate that this is due to the combination of the features within the F1 set, which resulted in better performance.

The second outcome in this paper is that data-driven models had higher accuracy in capturing the human-human co-manipulation. Hence, it is expected to be more accurate in the human-robot scenario. This points towards being correct since mathematical models require an accurate dynamical description of humans and robots to build such a system. Assuming that an accurate model is available of a given robot, building such a system will be feasible without a data-driven model. The counterargument to this is that, even if it is possible to obtain a very accurate mathematical model of the robotics system, this solution is not generic enough and can only be applied on one robot and one type of load. In other words, the model does not guarantee resilience to variations in the system.
The simulation results unexpectedly revealed that LR had the lowest displacement error, due to the combination of LR models with a feedforward control scheme as shown in Figure 7. Nonetheless, displacement accuracy is not the only aspect that needs to be considered. Once the human is physically in contact with the robot, the robot’s responses must not cause any injuries to the human. In other words, for movements that require high forces, jerkiness should be prohibited. Hence, even though LR models achieved the most accurate displacements, they had the highest interaction forces between humans and robots. This means that the human must do more work to co-manipulate an object with the robot. Hence, it will be more exhausting for the human to perform the required task.

8. Conclusions and Future Work

This paper utilised a data-driven approach to extract and model a human-human (demonstration) co-manipulation skill, which can be utilised to teach an industrial robot a human-like behaviour. The collected experimental data included F/T data, the manipulated object’s Cartesian position, and the EMG signal from the human muscles (follower). The collected data were then used to fit an RNN, an RF, an LR, and BT using four sets of features (F1, F2, F3, and F4). The accuracy of the fitted models was tested using unseen, randomly split test data, which illustrated that the sequential RNN trained on F1 features had the lowest RMSE compared to other ML models.

Moreover, this showed that the use of EMG sensory data positively impacted such a model’s accuracy. After that, the best data-driven model was compared with a simplified mathematical model. The results illustrated that the RNN outperformed the mathematical model; in fact, all data-driven approaches outperformed the mathematical models. Finally, the best data-driven models were validated in a simulated environment with an impedance controller to evaluate these approaches compared to MM and HM in a more realistic scenario. The results exhibited that data-driven models had higher accuracy and smoother trajectories in comparison with HM and MM.

In order to extend the work completed in this paper, it is essential first to outline the vision for the work. This project’s desired outcome was to create a controller that allowed co-manipulation tasks, including multiple robots following a human leader. However, substantial work needs to be completed until this vision can be established. The data-driven approach offers promising potential for more intuitive interfaces in HRC and, thus, more effective collaboration between humans and robots. Hence, testing the model experimentally in a human-robot manipulation scenario will be conducted in the future.

Author Contributions: Conceptualization, A.A.-Y., M.F. and A.B.; methodology, A.A.-Y. and M.F.; software, A.A.-Y.; validation, A.A.-Y., M.F. and A.B.; formal analysis, A.A.-Y.; investigation, A.A.-Y., M.F. and A.B.; resources, Intelligent Automation Centre lab; data curation, A.A.-Y., M.F.; writing—original draft preparation, A.A.-Y., M.F. and A.B.; writing—review and editing, A.A.-Y., M.F., A.B., T.B., P.F., E.-M.H. and N.L.; supervision, E.-M.H. and N.L.; project administration, E.-M.H. and N.L.; funding acquisition, E.-M.H. and N.L. All authors have read and agreed to the published version of the manuscript.

Funding: This project was funded by EPSRC:project (DigiTOP; EP/R032718/1).

Institutional Review Board Statement: The study was conducted according to the guidelines of the Declaration of Helsinki, and approved by the Ethics Approval Sub-Committee of Wolfson School, Loughborough University (25 November 2019).

Informed Consent Statement: Written informed consent has been obtained from the participant(s) to publish this paper.

Data Availability Statement: The data presented in this study are openly available in Loughborough University Repository at https://doi.org/10.17028/rd.lboro.12942122.v1.

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

1. Tsarouchi, P.; Makris, S.; Chryssolouris, G. Human-robot interaction review and challenges on task planning and programming. *Int. J. Comput. Integr. Manuf.* 2016, 29, 916–931. [CrossRef]

2. Kroemer, O.; Niekum, S.; Konidaris, G. A review of robot learning for manipulation: Challenges, representations, and algorithms. *arXiv* 2019, arXiv:1907.03146.

3. Zou, F. Standard for human-robot Collaboration and its Application Trend. *Aeronaut. Manuf. Technol.* 2016, 58–63, 76.

4. Heny, A.; Aouache, M.; Maoudj, A.; Akli, I. Human–robot interaction in industrial collaborative robotics: A literature review of the decade 2008–2017. *Adv. Robot.* 2019, 33, 764–799. [CrossRef]

5. ISO/TS 15066. *Robots and Robotic Devices Collaborative Robots*; ISO: Geneva, Switzerland, 2016.

6. Haddadin, S.; Haddadin, S.; Khoury, A.; Rokahr, T.; Parusel, S.; Burgkart, R.; Bicchi, A.; Albu-Schäffer, A. On making robots understand safety: Embedding injury knowledge into control. *Int. J. Robot. Res.* 2012, 31, 1578–1602. [CrossRef]

7. Kaneko, K.; Harada, K.; Kanehiro, F.; Miyamori, G.; Akachi, K. September. Humanoid robot HRP-3. In *Proceedings of the 2008 IEEE/RSJ International Conference on Intelligent Robots and Systems*, Nice, France, 22–26 September, 2008; pp. 2471–2478.

8. Robla-Gomez, S.; Becerra, V.M.; Llata, J.R.; Gonzalez-Sarabia, E.; Torre-Ferrero, C.; Perez-Oria, J. Working Together: A Review on Safe human-robot Collaboration in Industrial Environments. *IEEE Access.* 2017, 5, 26754–26773. [CrossRef]

9. Pratt, J.E.; Krupp, B.T.; Morse, C.J.; Collins, S.H. The RoboKnee: An exoskeleton for enhancing strength and endurance during walking. *In Proceedings of the Robotics and Automation (ICRA), 2004 IEEE International Conference, New Orleans, LA, USA, 26 April–1 May 2004*; Volume 3, pp. 2430–2435. [CrossRef]

10. Peternel, L.; Noda, T.; Petrič, T.; Ude, A.; Morimoto, J.; Babič, J. Adaptive control of exoskeleton robots for periodic assistive behaviours based on EMG feedback minimisation. *PLoS ONE* 2016, 11, e0148942. [CrossRef] [PubMed]

11. Dalle Mura, M.; Dini, G. Designing assembly lines with humans and collaborative robots: A genetic approach. *CIRP Ann.* 2019, 68, 1–4. [CrossRef]

12. Hayes, B.; Scassellati, B. Challenges in shared-environment human-robot collaboration. *Learning* 2013, 8, 1–6.

13. Oberc, H.; Prinz, C.; Głogowski, P.; Lemmerz, K.; Kuhlenkötter, B. Human Robot Interaction-learning how to integrate collaborative robots into manual assembly lines. *Procedia Manuf.* 2019, 31, 26–31. [CrossRef]

14. Schmidtler, J.; Knott, V.; Hözel, C.; Bengler, K. Human Centered Assistance Applications for the working environment of the future. *Occup. Ergon.* 2015, 12, 83–95. [CrossRef]

15. Weichhart, G.; Åkerman, M.; Akkaladevi, S.C.; Plasch, M.; Fast-Berglund, Å.; Pichler, A. Models for Interoperable Human Robot Collaboration. *IFAC-PapersOnLine* 2018, 51, 36–41. [CrossRef]

16. Yanco, H.A.; Drury, J. October. Classifying human-robot interaction: An updated taxonomy. In *Proceedings of the 2004 IEEE International Conference on Systems, Man and Cybernetics* (IEEE Cat. No. 04CH37583), The Hague, The Netherlands, 10–13 October 2004; Volume 3, pp. 2841–2846.

17. Sieber, D.; Deroo, F.; Hirche, S. November. Formation-based approach for multi-robot cooperative manipulation based on optimal control design. In *Proceedings of the 2013 IEEE/RSJ International Conference on Intelligent Robots and Systems*, Tokyo, Japan, 3–7 November 2013; pp. 5227–5233.

18. Ikeura, R.; Inooka, H. Variable impedance control of a robot for cooperation with a human. *Proc. IEEE Int. Conf. Robot. Autom.* 1995, 3, 3097–3102.

19. Ikeura, R.; Morita, A.; Mizutani, K. Variable damping characteristics in carrying an object by two humans. In *Proceedings of the 6th IEEE International Workshop on Robot and Human Communication. RO-MAN’97 SENDAI*, Sendai, Japan, 29 September–1 October 1997.

20. Rahman, M.M.; Ikeura, R.; Mizutani, K. Control characteristics of two humans in cooperative task and its application to robot Control. *IECON Proc. Ind. Electron. Conf.* 2000, 1, 1773–1778. [CrossRef]

21. Bussy, A.; Gergondet, P.; Kheddar, A.; Keith, F.; Crosnier, A. Proactive behavior of a humanoid robot in a haptic transportation task with a human partner. In *Proceedings of the 2012 IEEE RO-MAN: The 21st IEEE International Symposium on Robot and Human Interactive Communication*, Paris, France, 9–13 September 2012.

22. Galin, R.; Meshcheryakov, R. Review on human-robot Interaction During Collaboration in a Shared Workspace. In *International Conference on Interactive Collaborative Robotics*; Springer: Cham, Switzerland, 2019; pp. 63–74.

23. Huang, J.; Pham, D.T.; Wang, Y.; Qu, M.; Ji, C.; Su, S.; Xu, W.; Liu, Q.; Zhou, Z. A case study in human–robot collaboration in the disassembly of press-fitted components. *Proc. Inst. Mech. Eng. B J. Eng. Manuf.* 2020, 234, 654–664. [CrossRef]

24. Hakonen, M.; Piitulainen, H.; Visala, A. Current state of digital signal processing in myoelectric interfaces and related applications. *Biomed. Signal Process. Control.* 2015, 18, 334–359. [CrossRef]

25. Mielke, E.; Townsend, E.; Wingate, D.; Killpack, M.D. human-robot co-manipulation of extended objects: Data-driven models and control from analysis of human-human dyads. *arXiv* 2020, arXiv:2001.00991.

26. Townsend, E.C.; Mielke, E.A.; Wingate, D.; Killpack, M.D. Estimating Human Intent for Physical human-robot Co-Manipulation. *arXiv* 2017, arXiv:1705.10851.

27. Buerkle, A.; Lohse, N.; Ferreira, P. Towards Symbiotic human-robot Collaboration: Human Movement Intention Recognition with an EEG. In *Proceedings of the UK-RAS19 Conference: “Embedded Intelligence: Enabling & Supporting RAS Technologies* Proceedings, Loughborough, UK, 27 January 2019; pp. 52–55.
28. Wang, P.; Liu, H.; Wang, L.; Gao, R.X. Deep learning-based human motion recognition for predictive context-aware human-robot collaboration. CIRP Ann. 2018, 67, 17–20. [CrossRef]
29. Mukhopadhyay, S.C. Wearable sensors for human activity monitoring: A review. IEEE Sens. J. 2014, 15, 1321–1330. [CrossRef]
30. Peternel, L.; Fang, C.; Tsagarakis, N.; Ajoudani, A. A selective muscle fatigue management approach to ergonomic human-robot co-manipulation. Robot. Comput. Integr. Manuf. 2019, 58, 69–79. [CrossRef]
31. Chowdhury, R.H.; Reaz, M.B.; Ali, M.A.B.M.; Bakar, A.A.; Chellappan, K.; Chang, T.G. Surface electromyography signal processing and classification techniques. Sensors 2013, 13, 12431–12466 [CrossRef] [PubMed]
32. Lazar, J.; Feng, J.H.; Hochheiser, H. Measuring the human. In Research Methods in Human-Computer Interaction, 2nd ed.; Hochheiser, E., Ed.; Morgan Kaufmann: Boston, MA, USA, 2017; pp. 369–409.
33. Géron, A. Hands-on Machine Learning with Scikit-Learn, Keras, and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems; O’Reilly Media: Sebastopol, CA, USA, 2019.
34. Naufel, S.; Glaser, J.I.; Kording, K.P.; Perreault, E.J.; Miller, L.E. A muscle-activity-dependent gain between motor cortex and EMG. J. Neurophysiol. 2019, 121, 61–73. [CrossRef] [PubMed]
35. Peternel, L.; Tsagarakis, N.; Ajoudani, A. A human–robot co-manipulation approach based on human sensorimotor information. IEEE Trans. Neural Syst. Rehabil. Eng. 2017, 25, 811–822. [CrossRef] [PubMed]
36. Buerkle, A.; Al-Yacoub, A.; Ferreira, P. An Incremental Learning Approach for Physical human-robot Collaboration. In Proceedings of the 3rd UK Robotics and Autonomous Systems Conference (UKRAS 2020), Lincoln, UK, 17 April 2020.
37. Buerkle, A.; Eaton, W.; Lohse, N.; Bamber, T.; Ferreira, P. EEG based arm movement intention recognition towards enhanced safety in symbiotic Human-Robot Collaboration. Robot. Comput. Integr. Manuf. 2021, 70, 102137. [CrossRef]
38. Li, Y.; Ge, S.S. Human–robot collaboration based on motion intention estimation. IEEE/ASME Trans. Mech. 2013, 19, 1007–1014. [CrossRef]
39. Herbrich, R.; Graepel, T.; Obermayer, K. Support vector learning for ordinal regression. In IET Conference Proceedings; Institution of Engineering and Technology: London, UK, 1999
40. Lieber, R.L.; Jacobson, M.D.; Fazeli, B.M.; Abrams, R.A.; Botte, M.J. Architecture of selected muscles of the arm and forearm: Anatomy and implications for tendon transfer. J. Hand Surg. 1992, 17, 787–798. [CrossRef]
41. ATI Industrial Automation. Force/Torque Sensor Delta Datasheet. Available online: https://www.ati-ia.com/products/ft/ft_models.aspx?id=Delta (accessed on 18 February 2020).
42. Myoware. EMG Sensor Datasheet. Available online: https://cdn.sparkfun.com/datasheets/Sensors/Biometric/MyowareUserManualAT-04-001.pdf (accessed on 11 February 2020).
43. Al-Yacoub, A.; Buerkle, A.; Flanagan, M.; Ferreira, P.; Hubbard, E.M.; Lohse, N. Effective human-robot collaboration through wearable sensors. In Proceedings of the 2020 25th IEEE International Conference on Emerging Technologies and Factory Automation (ETFA), Cranfield, UK, 3–4 November 2020; Volume 1, pp. 651–658.
44. Brunton, S.L.; Kutz, J.N. Data-Driven Science and Engineering: Machine Learning, Dynamical Systems, and Control; Cambridge University Press: Cambridge, UK, 2019.
45. He, T.; Xie, C.; Liu, Q.; Guan, S.; Liu, G. Evaluation and comparison of random forest and A-LSTM networks for large-scale winter wheat identification. Remote Sens. 2019, 11, 1665. [CrossRef]
46. Reinhart, R.F.; Shareef, Z.; Steil, J.J. Hybrid analytical and data-driven modeling for feed-forward robot control. Sensors 2017, 17, 311. [CrossRef]
47. Naik, G.R. Computational Intelligence in Electromyography Analysis: A Perspective on Current Applications and Future Challenges; InTech: Rijeka, Croatia, 2012.
48. Wyeld, T.; Hobbs, D. Visualising Human Motion: A First Principles Approach using Vicon data in Maya. In Proceedings of the 2016 20th International Conference Information Visualisation (IV), Lisbon, Portugal, 19–22 July 2016; pp. 216–222.
49. Al-Yacoub, A.; Zhao, Y.C.; Eaton, W.; Goh, Y.M.; Lohse, N. Improving human robot collaboration through Force/Torque based learning for object manipulation. Robot. Comput. Integr. Manuf. 2021, 69, 102111. [CrossRef]