Hand-in-Hand: Investigating Mechanical Tracking for User Identification in Cobot Interaction

Alia Saad
alia.saad@uni-due
University of Duisburg-Essen
Essen, Germany

Max Pascher
max.pascher@uni-due.de
University of Duisburg-Essen
Essen, Germany

Khaled Kassem
khaled.k.kassem@tuwien.ac.at
TU Wien
Vienna, Austria

Roman Heger
roman.heger@uni-due.de
University of Duisburg-Essen
Essen, Germany

Jonathan Liebers
jonathan.liebers@uni-due.de
University of Duisburg-Essen
Essen, Germany

Stefan Schneegass
stefan.schneegass@uni-due.de
University of Duisburg-Essen
Essen, Germany

Uwe Gruenefeld
uwe.gruenefeld@uni-due.de
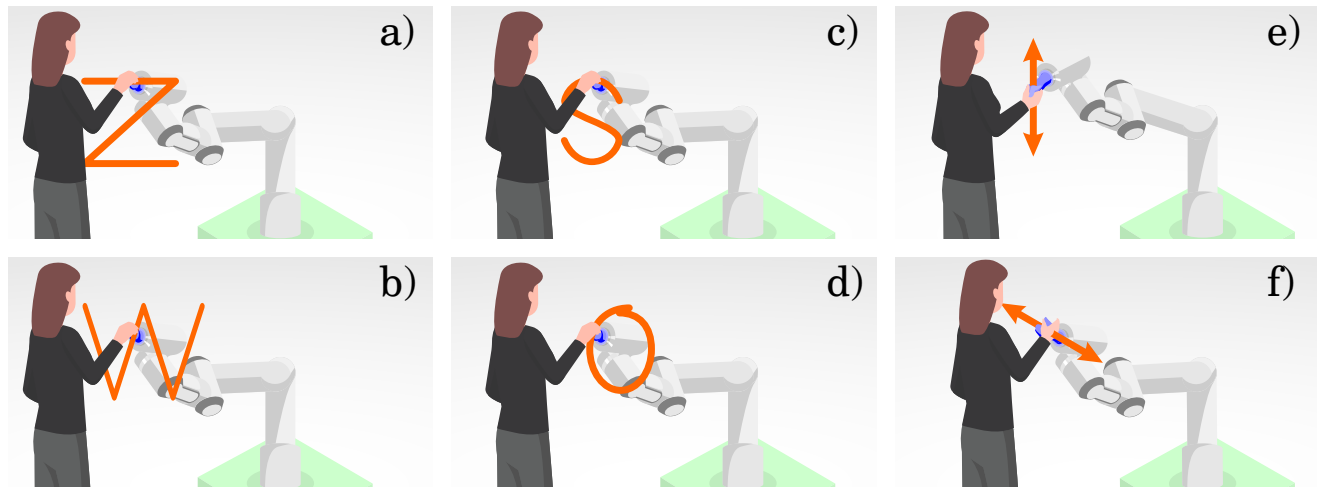University of Duisburg-Essen
Essen, Germany

Figure 1: In this work, we examine performing direct interactions with the robotic arm as a behavioral biometric. We investigated six different gestures for three movement categories: Circular: (a) and (b), Linear (c) and (d), and Human gestures: formal handshake (e) and informal handshake (f). Using Random Forest classifiers, we achieved an average F1-scores up to 0.87.

ABSTRACT

Robots play a vital role in modern automation, with applications in manufacturing and healthcare. Collaborative robots integrate human and robot movements. Therefore, it is essential to ensure that interactions involve qualified, and thus identified, individuals. This study delves into a new approach: identifying individuals through robot arm movements. Different from previous methods, users guide the robot, and the robot senses the movements via joint sensors. We asked 18 participants to perform six gestures, revealing the potential use as unique behavioral traits or biometrics, achieving F1-score up to 0.87, which suggests direct robot interactions as a promising avenue for implicit and explicit user identification.

CCS CONCEPTS

• **Human-centered computing** → **Empirical studies in HCI**; **Laboratory experiments**; **User studies**; • **Computer systems organization** → **Robotics**; **Robotics**.

KEYWORDS

behavioral biometrics, cobots, human-robot interaction, human-robot collaboration

1 INTRODUCTION

The usage of robots is growing steadily beyond industrial applications, with robots covering a wide range of functions, from simple domestic tasks (e.g., vacuum cleaning [50]) to complex technical assignments (e.g., product assembly [14]). Not only did their applications diversify, robots – in this case, frequently referred to as *cobots* – have become increasingly collaborative and often work closely together with humans [41]. One notable example is the increased user independence achievable through domestic care cobots when supporting people with disabilities in their everyday tasks [12, 42]. Building on this, a particular subset of cobots – i.e., robotic arms – serve as an extension of human capabilities, creating a symbiosis that can help users achieve their goals [36].

One important particularity of cobots is that they are frequently shared between users. This can be on the operator level, e.g., cobots being used by and shared between different workers across shifts [46]. Alternatively, on a recipient level, medical caregivers might use cobots to care for several patients, often a necessity given the high cost and limited availability of assistive technologies [42]. While sharing cobots is often the most efficient resource usage, it poses an additional challenge for developers. There are several use cases where human ergonomics are prioritized to ensure a safe, comfortable, and efficient collaboration. For instance, in rehabilitation, where robots need to adjust to a patient's range of motion [35]. Another example is workbench assistance, where cobots are working with humans on tasks that require precision, repeatability, or manual dexterity such as assembly assistance or handling parts [23]. In these scenarios, effective robotic solutions must be supported with identification-based personalization options to accommodate different user preferences and – most importantly – consider ergonomic concerns (a core part of worker well-being and Industry 5.0 [28]). As Völkel et al. points out, personalization is not only a pressing requirement for customization but also for elevating acceptance and trust towards robots [53]. Hence, including functions that support identifying and distinguishing between users is essential for a personalized and safe collaboration.

Current solutions range from knowledge-based methods like passwords to physical tokens as identifiers. However, these solutions are not implicit, requiring additional action from the user [2]. Moreover, they are liable – physically or via observation attacks – to be forgotten, lost, or stolen. *Behavioral biometrics* utilize user interaction with a system for implicit identification, negating many current approaches' drawbacks. This strategy has been successfully applied in various contexts, including mobile phones [31] and virtual reality [26, 44]. Behavioral biometrics has yet to be analyzed in detail for interacting with cobots. Huang et al. suggests that a robotic arm mimicking users' movements contains similar behavioral features that can be used as a distinct trait for identification [19]. However, whether cobots can successfully interpret direct users' movements for unambiguous identification still needs to be explored.

This work investigates how direct interactions with a robotic arm can use sensor data to enable user identification. In a laboratory study (N=18), we compared the feasibility of six different gestures and their resulting identification accuracy. Our results show that user identification can work across gestures, with a mean identification F1-score of 0.87. To validate our method, we conducted a second round of data collection from a subset of users from the first session (N=6). Although results showed a decrease in performance with a mean F1-score accuracy of 0.54, they reflect a normal performance of behavioral identification over multiple sessions. We demonstrate that accurate biometrics-based user identification is possible without deploying additional sensors. These findings imply that existing systems can be easily extended with user identification capabilities. Moreover, we expect this approach to serve as a continuous and implicit identification method, enabling cobots to consistently verify the user's identity while performing other tasks. Here, we summarize our work's contributions as an empirical evaluation of user identification performance across and between different gestures directly performed with and mechanically tracked by a robot arm.

2 RELATED WORK

Our work intersects two research fields: Human-Robot Interaction (HRI) – particularly collaborative robots – and behavioral biometrics for user identification.

Human-Robot Interaction. Collaborative robots – cobots – are becoming increasingly ubiquitous, going beyond manufacturing industries to other fields like domestic care [5, 40]. As such, the research field of Human-Robot Collaboration (HRC) is diversifying. It encompasses categories based on levels of environment sharing [33, 45, 51], type of cooperation [29], and interaction [4, 17]. Focusing specifically on close-proximity interactions, previous work has explored how robots adapt to human movements and behavior, such as maintaining appropriate distance [37], or predicting and avoiding collision [18]. Other studies underline the collaboration effect by analyzing how robot movements adapt to specific human needs, including adjusting to user fatigue [43], providing personalized assistance based on user skill level [7], or adjusting the robot movements to maximize human comfort during collaboration [10]. According to Bonci et al. [8], more efficient HRC can be achieved when appropriate techniques and sensors are used.

Studies have shown that physical proximity to a cobot during collaboration shows no adverse effects when safety elements are fulfilled [13]. This agrees with findings by Maurtua et al. [30], where 97% of participants in a responsive collaborative setup study declared that the investigated interaction type is foreseeable to become prevalent. Further, human embodiment can positively affect both the perception of and the trust in the robot [56]. Previous work has shown that embodied touch-based interactions with robots can increase the non-verbal communication capabilities of the robot, as well as attenuate the human's stress responses [55]. Interactive perception, or the combination of physical and traditional perception methods, has expanded its range of applications. However, the view could be occluded by an arm in motion, affecting the perception of

other sensors [25]. On the other hand, kinesthetic teaching – *Learning from Demonstration* –is a well-established method used to teach robots new skills without the need for robotics or programming expertise from the user, with a high level of trust [6, 52, 58].

Behavioral Biometrics. Behavioral biometrics relies on recognizing user behavioral patterns for identification and authentication [3]. This approach involves analyzing various aspects of behavior, such as the characteristics of handwriting [32], the timing of keystrokes [57], distinct walking patterns (gait) [54], speech patterns [38], stylus usage [9], and other behavioral features that collectively define an individual's behavior such as tracked body movements in VR [26]. Knowing the user's identity currently interacting with a system allows for improving the interaction through personalizations, e.g., customization of the interface or loading preferred settings [22]. While standard identification approaches require explicit input, behavioral biometrics offers an implicit alternative. Different types of behavior have been used as biometric features, such as arm and head movements – both well-suited biometric features – in virtual reality context [26, 44]. Similarly, hand input on standard interface elements in virtual and augmented reality also allows for accurate user identification [27]. Going one step further, Pohl et al. uses a single button press on a physical button to distinguish users [47]. Their approach recognizes specific users through a single press, using a combination of springs and sensors. Smartphones also use single button interactions for accessible PIN code entry [24]. These techniques highlight the importance and potential of broadening identification strategies beyond conventional methods. Exploring behavioral biometrics based on gestures is a recently emerging trend. Imura and Hosobe used a three-dimensional hand gesture-based method, allowing its user to move the user's hand without touching an input device [21]. 2D gestures using a smart pen showed an identification rate of 87% in a study by Schrapel et al. [49]. In mid-air, various researchers explored how 3D signatures can be used as a biometric [11, 48]. Similarly, Huang et al. explore midair gestures in combination with a robotic arm. They use an optical tracking system tracking the hand and a robotic arm. The robotic arm mimics the user's behavior, and they found that the robotic arm's movement contains behavior that can be used as a biometric variable [19, 20].

Summary. Past research emphasizes the wide variety of cobots applications, and the essential need to identify the collaborating users. Additionally, there is a growing focus on behavioral biometrics, given its resilience to attacks and its suitability for both implicit and explicit identification.

3 IDENTIFYING USERS USING A ROBOTIC ARM

It is clear that behavioral biometrics rely not only on users' behaviors but also on physiological traits that play a vital role. Here, we focus on the specific case of direct mechanical manipulation of a robotic arm, as it limits the range of gestures, and accordingly, poses restrictions on gesture execution.

3.1 Approach

This work proposes a novel approach for accurate user identification using a robotic arm. In particular, we focus on gestures as input techniques through direct manipulation of robot end-effectors. Such manipulation can happen while the user moves the robot to teach a specific task or movement, or while doing gestures assigned to a specific command that the robot interprets. The manipulation of the robotic arm can be measured by using the mechanical tracking of the robot arm joints. Thus, this approach does not require additional tracking (e.g., optical tracking).

To examine the feasibility of this approach, we specify two categories of gestures: **letter-based gestures** and **handshake gestures**. The former was selected from the work of Huang et al. [20] and could be used to enter specific commands. Four letters in total were selected, two simulating circular movements (*"O"* and *"S"*), and two for linear movements (*"W"* and *"Z"*). Secondly, we consider a human-like interaction with the robot, where a handshake is performed. Here, we specified two types of handshakes (*"formal"* and *"informal"*), simulating vertical and horizontal handshake interactions. This could be used to start or during the interaction with an anthropomorphic end-effector [39].

3.2 Study Design

We conducted a within-subject controlled laboratory study across two sessions. During each session, we measured the variable (*gesture*) with six levels (four letter-based and two handshake gestures – see Figure 1). Our goal was to elicit user input for a range of gestures, showing the generalizability of our approach rather than a comparison of different gestures. As the two types of gestures (letter and handshakes) require different movement inputs, we used two distinct grasp handles as the end-effector of the robot arm (Figure 3). This also considers that robots are typically used for a range of tasks that can involve different end-effectors. For letter-based gestures, we used a round knob that allows for easy manipulation of the robot arm in 3D space. As for gestures mimicking greetings, we used a model resembling a human hand. For each of our six conditions, participants were asked to repeat each gesture five times, resulting in 30 gestures per session. We counterbalanced all conditions in both sessions using a complete Latin-square design. Our dependent variables were (i) *classification accuracy* measured with the F1-score of the trained classifiers, (ii) task load measured with the raw NASA-TLX [15, 16], and (iii) *subjective feedback* consisting of individual ratings for the most liked and disliked gestures. We also asked participants open-ended questions to get further qualitative insights. All dependent variables were measured in the first session, while we looked at classification accuracy across sessions. Our work is motivated by the research question **to what extent users can be identified from robot motion when they directly interact with the robot's end-effector through different gestures?** Inspired by related work, we derived these three hypotheses:

- **H₁** Users are identifiable with high accuracy through their behavior contained in direct gestures with a robot arm.
- **H₂** We expect user identification accuracy to drop when training and testing with different sessions.
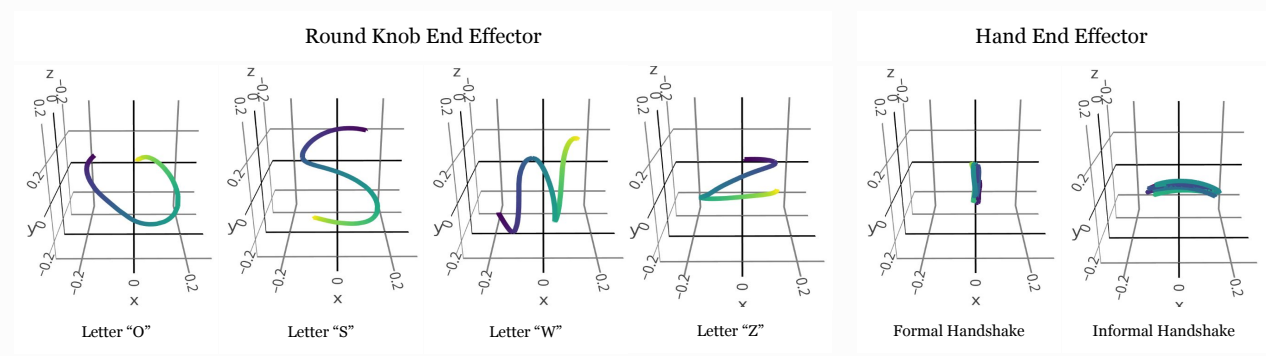
Figure 2: Example plots of robot end-effector position in 3D space (over time) for all six gestures. All 3D scatter plots show one repetition from one randomly selected study participant to illustrate the captured temporal-spatial data. Blue: start, yellow: end of movement.
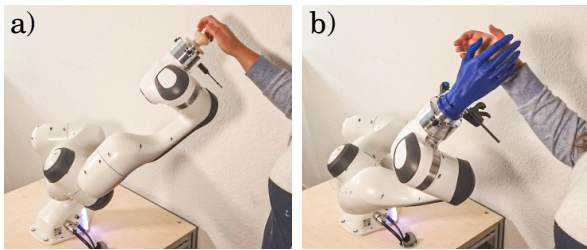

Figure 3: Study setup showing the interaction with the robotic arm using a knob (a) and a hand (b).

H₃ We do not expect any significant differences in temporal, mental or physical demands between gestures.

3.3 Participants and Procedure

We invited 18 participants (male = 11, female = 7) via mailing lists and social networks. Their ages varied between 22 and 55 years ($M = 29.44$, $SD = 7.42$). Except for one person, all participants were right-handed. However, the left-handed person was trained to write using their right hand at a young age and faced no difficulty using their non-dominant hand. As this work focuses on direct interaction with the robot, no forms of motor function limitations or injuries were reported. We wanted to explore the impact of time on user behavior. To that end, after nine months, we conducted a second session with a subset of six previous participants. All participants were compensated for their participation with 10 Euros. We received ethical clearance from our local ethics committee (#20220721).

Upon arrival at our lab, we informed the participants about the purpose of the study. Next, they provided informed consent and filled out a demographic form (in the first session). One participant had prior experience working with cobots. However, none of the participants had prior knowledge of direct mechanical interaction with the robotic arm. To that end, the study experimenter proceeded to demonstrate the movements to the participants and allowed them to move the robot. This step was done for each gesture. To consistently evaluate all participants, the robot was automatically put

in the same initial position before every interaction. We ensured that the participants were not within the robotic arm range while performing this step. Except for the handshake, which had an initial point of 110 cm, all other gestures had a starting point at a height of 130 cm. The collection setup also involves putting the robot in the safe interaction or guiding mode. We placed a clamp to simulate a user pressing the robot's free guiding mode enabling buttons, allowing a free movement of all seven joints of the robot. The clamp was carefully placed to hold the buttons, and when moved, the robot automatically stops, as the free guiding mode is disabled. In our setup, the robot only automatically moves when returned to the initial position, otherwise, it is controlled by the participants. Therefore, the safety of the participants was preserved throughout the entire study procedure. Once notified, the participant proceeded to apply the required interaction till they were notified to stop. Every gesture is recorded 5 times in the first and 7 times in the second session. After the first session, participants answered a NASA Task Load index (TLX) questionnaire [16], following each gesture condition, to potentially identify differences between the gestures. Moreover, once all tasks were done in the first session, participants filled out a post-study questionnaire. The survey collected participants' subjective feedback on their preferred and least preferred gestures, along with the reasons for their choices and suggestions for additional gestures. These insights, comparing between hand and knob handle, inform the development of a gesture vocabulary tailored to the robotic arm's movement constraints.

3.4 Apparatus

We conducted the study in a spacious room in our lab (4m x 8m) with a brightness of 200 Lux. We used a Franka Emika Panda robot in this study[1]. The robot was mounted on a table (height = 61 cm). For the end-effectors, we 3D printed a hand and knob model that we acquired from an open source library[2]. Both end-effectors were printed with PLA. We did not mount any additional sensors but solely relied on the sensor data provided by the robot arm itself.

---

[1]https://frankaemika.github.io/docs/overview.html, last accessed: October 20, 2023
[2]https://www.thingiverse.com, last accessed: October 20, 2023

3.5 Data Analysis

The goal of this work is to identify users based on their direct movements with the robotic arm. We used the data from the first session for identification across all 18 participants, whereas we used the second session data to assess the identification performance of the participants subset (N=6). We acquired the state of all the robot's joints (7 in total) in terms of position, velocity, and effort. The *joints* space limits were collected in relative radian values (*rad*). Moreover, we validated the recorded data by replaying it in Unity3D[3], converting the *end-effector* positions to the absolute position and rotations values in X, Y, and Z axes, as seen in the example in Figure 2. After visual inspection of all data, we started the preprocessing by extracting the performed gestures. Therefore, idle states before and after the interaction with the robot were omitted, ensuring that only the user's movements during gesture performance were used in classification. For training, we computed different feature sets. We had two feature sets using descriptive statistical values (min, max, mean, and standard deviation) and two feature sets using only the first row, meaning the first row of the performed gesture execution. Within each of these pairs of feature sets, we first look into all three aforementioned values per joint for all 7 joints, resulting in a set of 84 features per recording (7 joints × 3 joint states × 4 descriptive features) and then look into all values from the replay (position and rotation of end-effector), resulting in a set of 24 features per recording (2 end-effector states × 3 spatial coordinates (X,Y,Z) × 4 descriptive features). Each of the four resulting feature sets is then tested within the first session using 5-fold cross-validation with 80/20 splits and also tested between sessions using 2-fold cross-validation (training with one session and testing with the other session). The latter, we did as recent publications highlight behavioral changes over time have an impact on behavioral biometrics [26, 34]. For each fold, we trained a Random Forest (RF) classifier (*n_estimators* = 300). Given the limited size of our dataset, we opted for a straightforward train/test split and decided against parameter hypertuning. Therefore, we used the default values for all algorithms, except the number of trees of the random forest classifier was set to 300 from our previous experience developing behavioral biometric systems. We applied RF because it is not a black box model and allows us to look at the importance of features, contributing to the explainability of our approach. Finally, we use the resulting F1-scores of our classifiers for inferential statistics. Here, we applied non-parametric tests, as we do not assume the normality of our data. Though our experiment has less power, our results are more robust and not influenced by outliers or deviations from specific distributions.

4 RESULTS

Here, we present the findings from our user study and training of classifiers. The objective measures relate to the performance of the classifiers (trained with the joints and end-effector data), while the subjective measures are metrics gathered during the user study.

4.1 Objective Measures

We showcase the results derived from the interactions, focusing on the descriptive features of the robot's *joints* and the calculated *end-effector* values. In our exploration of user identification for distinct gestures executed directly with the robot, the first approach we took is the classical descriptive metrics for these joint and end-effector values, including the minimum, maximum, mean, and standard deviation. We followed a 5-fold cross-validation, where every repetition was treated as a test set, and the other repetitions of each gesture per participant were used for training. The k-fold cross-validation (k=5) results for the *joints* values reached an average F1-score of 0.87 (Md = 0.9, IQR = 0.06), whereas a similar cross-validation for the calculated *end-effector* yielded results of a mean F1-score of 0.80 (Md = 0.82, IQR = 0.1). Both are plotted in the Figure 4 in the left plot. We sorted the features based on their importance, i.e., contribution to the classification accuracy. The list of the ten most important features in the joints dataset contained six position features of different joints. In regard to the end-effector data, we found that position and rotation values from the Y and Z axes were more important to the classifier in comparison to the values in the X axis.

To understand the performance of the different gestures, we investigated user identification accuracy for each gesture individually. We used k-fold cross-validation (k=5) to train five classifiers for each gesture. As every participant performed each gesture five times, each repetition served as our test set once, while the other four repetitions were used for training. Thereafter, we aggregated the F1-scores of each fold for each participant individually by calculating the mean across all five folds. The mean (median; interquartile range) F1-scores over all participants for each gesture are (in descending order): Letter Z=0.91 (Md=0.93; IQR=0.20), Letter W=0.87 (Md=0.93; IQR=0.20), Informal Handshake=0.87 (Md=0.93; IQR=0.20), Letter S=0.86 (Md=0.87; IQR=0.25), LetterO=0.85 (Md=0.93; IQR=0.23), and Formal Handshake=0.78 (Md=0.80; IQR=0.26), as plotted in Figure 4 (right). A Shapiro-Wilk-Test showed that our data is not normally distributed (p < 0.001), and therefore, we applied non-parametric tests. Post-hoc tests using a Wilcoxon test with Bonferroni correction did not show specific pairs of gestures with significant differences in F1-scores.

Following the promising results of the descriptive features, we investigated the extent to which the beginning of the gesture execution contributes to the classification accuracy. Although we standardized the robot's initial static pose for all participants, our emphasis was on the pose at which participants started to perform the gesture (these differ between participants as we extracted the execution of the gesture from the complete recording of the repetition). Accordingly, we trained our classifier with only the first row of each repetition. Following the same approach of 5-fold cross-validation, the results showed an average F1-score of 0.58 (Md = 0.56, IQR = 0.04) and 0.37 (Md = 0.39, IQR = 0.01) for the *joints* and *end-effector* data, respectively.

Examination of Behavior Across Extended Periods. Lastly, we evaluated the behavior variability across extended periods through a between-session analysis, where both sessions were used for training and testing. Compared to our previous analysis, we could only

---

[3]https://github.com/niccarey/FrankaPanda_Unity, last accessed: October 20, 2023
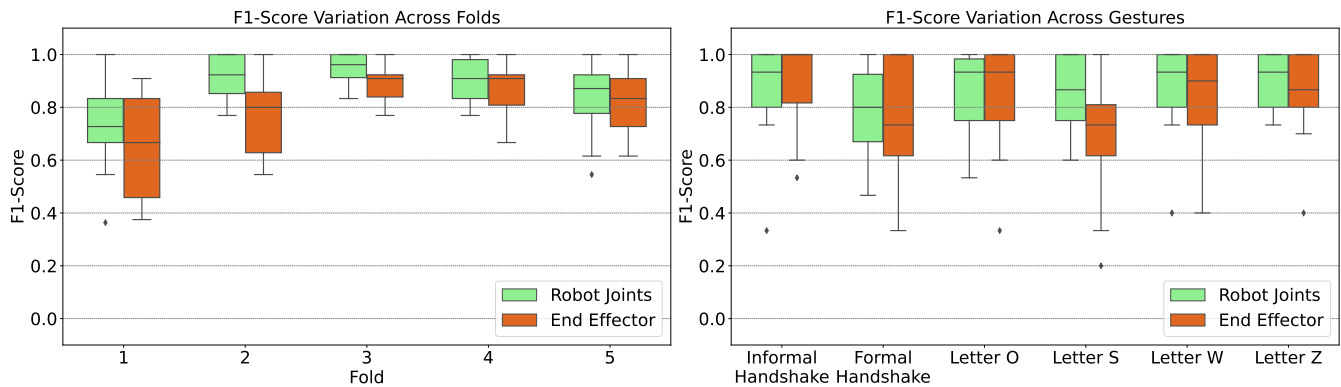
Figure 4: Boxplots of F1-scores for user identification 5-fold cross-validation for all 6 gestures for descriptive features(left), and between gestures (right). Both boxplots show the results in terms of robot joints and end-effector values.

differentiate between six participants, as we were unable to re-recruit more of the initial set of participants for the second session. The results showed an average F1-score of 0.54 and 0.51, for the *joints* and *end-effector* data, respectively.

4.2 Subjective Measures

We gathered ordinal data from our Likert items asked during the user study. Accordingly, we directly apply non-parametric tests for inferential statistics. For task load, we asked participants to fill out a raw NASA-TLX questionnaire after each condition [15, 16]. The median (interquartile range) task load scores for each gesture are (in ascending order): *LO*=10.83 (IQR=11.04), *LZ*=10.83 (IQR=11.25), *Formal Handshake*=10.83 (IQR=15.83), *Informal Handshake*=12.08 (IQR=7.92), *LW*=12.92 (IQR=8.13), and *LS*=14.58 (IQR= 13.75). To compare task load between gestures, we performed a Friedman test that did not reveal any significant differences between the conditions (χ^2(5)=1.48, p=0.915, N=18). For individual dimensions of the raw NASA-TLX, refer to Figure 5 for the task load scores.

After all conditions, we asked participants which gesture they liked most and which one they liked the least. Concerning the most liked gesture, participants stated that they liked the vertical handshake *(Formal Handshake)* the most (n=9), followed by the letter "O" *(LO)* (n=4), the horizontal *(Informal Handshake)* (n=3), and letter "S" *(LS)* and letter "W" *LW* with each one vote. With regard to the least liked gestures, participants mentioned that they disliked the letter "Z" *(LZ)* the most (n=6), followed by the *(Informal Handshake)* (n=5), and then voted equally often for the letters "O", "S", "W" (n=2). One participant voted for the formal handshake as the least liked one. When asked about the reason for selecting the most liked gesture, 7 out of the 9 participants choosing handshake stated that it was natural and "easy to perform", while those who chose the letter "O" *(LO)* stated that it provides an "intuitive" sense for the interaction, and "easier to perform". On the other hand, participants said that sharp lines were difficult to execute, that the sharpness made it "mote demanding to think about and to execute". Finally, we asked the participants to create their own gestures. Responses included "fist bumps", other letters or symbols such as *I, L*, or *the interrogation mark (?)*.

5 DISCUSSION

Mechanical Tracking Enables Behavioral Biometrics. We investigated user identification through mechanical tracking using robot arm joints. By selecting six different gestures unified across participants, we found that movements encoded in the joints could, in fact, be used as a behavioral metric for user identification. Following this approach, we were able to leverage the usage of the seven joints of the robotic arm sensors to gain information about the identity of the user beyond the typical use cases (e.g., manufacturing). With a base chance of 1/16 (6.25%) and our F1-score accuracy reaching up to 0.95 for some folds, we accept our first hypothesis (**H₁**) that *users are identifiable through the behavior contained in direct gestures with a robot arm with a high accuracy*. The results show a minor decrease in performance when compared to the related work of Huang et al., who reported 100% accuracy using an optical tracking system, yet with a lower number of participants (N=10) and more repetitions (40) using similar letter-based gestured [19, 20]. We assume that – besides the number of participants and sample size – employing mechanical tracking results in a lower accuracy since it reduces some of the variance in the movement through its inertia. However, in our approach, we do not require additional tracking hardware, allowing each robot to be used without modifications for user identification, highlighting the impact of our study findings.

Importance of Individual Features. Interestingly, we found that the starting point of the interaction plays an important role in the user identification performance. When looking at our results for the "first-row" feature sets, we reached a classification accuracy of up to 0.69 for one of the folds (which is only 20% less than the feature sets with all descriptive features reached). This means that we could infer that the starting point of the interaction contributes greatly to the user identification. We believe this might result from the diversity in physiological characteristics of the users, such as arm length and body height. In previous work, it has been demonstrated that physiological traits are often included in the elicited behavioral data [26]. As many gestures share the same starting point (e.g., the letter "Z" and "S"), it points towards the generalizability of our findings, suggesting that our trained classifiers may also scale
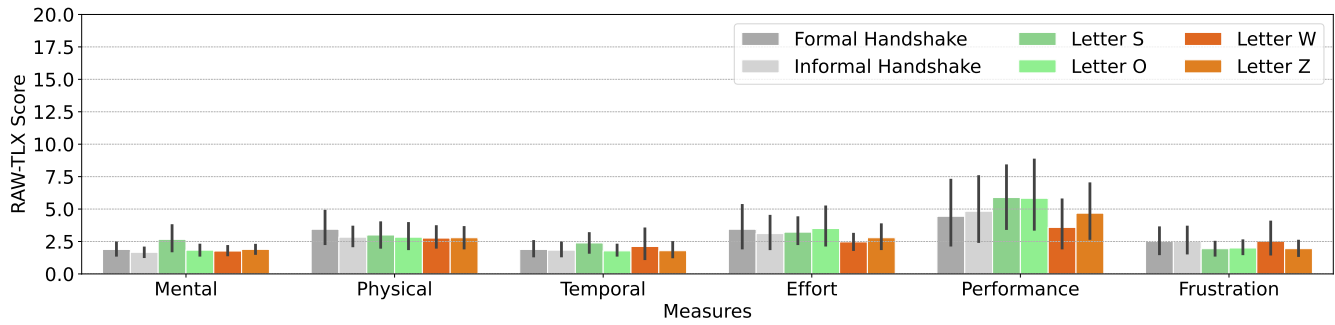
Figure 5: Raw NASA-TLX medians in six subscales across the six gestures. Scores range from 0 (low demand/perfect) to 20 (high demand/failure). The results show that no gesture was demanding in any sub-scale.

towards gestures not seen during training (if the utilize the same starting point). However, more research is needed to confirm this.

Gesture Performance Across Sessions. With a second session of collection, our findings show an observable drop in accuracy performance despite the smaller sample size (18 participants in the first and six participants in the second session). However, the resulting accuracy is not only in line with our second hypothesis (H_2) that expects an accuracy performance change, but also with the accuracy declines typically found in papers using hand movements and behavioral biometrics (e.g., Liebers et al. [26]; Pfeuffer et al. [44], and Miller et al. [34]). As a consequence, the behavioral biometric system needs continuous training to remain of good quality. While the drop in accuracy is quite drastic in our case, we believe this is due to several factors. First, the time difference between the two sessions is large (9 months), which could result in imperceptible yet considerable physiological changes affecting the identification performance. In addition, due to the unfamiliarity of the direct robot interaction, participants might exhibit inconsistent of fluctuating behavior. Therefore, we expect improved and more robust identification performance with continuous usage of this approach.

Direct Mechanical Interactions Are Not Demanding. As we expected in (H_3), such gesture-based interactions would not be demanding to the users, particularly in terms of temporal, mental, or physical load. We based our hypothesis on the nature of our approach, which is not time stressing, as there was no temporal condition to complete the task. Moreover, we considered in our design gestures that are already known to the users (letters or handshakes). Performing such gestures would not require prior training or induce a particular mental load to perform the tasks. Additionally, we did not specify any requirements to perform the gestures (e.g., certain dimensions of the letters) to avoid altering their behaviors. Subjective feedback findings supported these arguments. Participants' NASA-TLX responses show that the tasks were not mentally, physically, or temporally demanding. In the open questions, no participant reported any issues with performing any of the tasks. They rated their most and least liked gestures based on preference and convenience.

Limitations and Future Work. In our user study, we used two distinct robot end-effectors – knob and hand – to accommodate the

different types of gestures that we investigated. The human-like hand allowed for natural greetings, while the knob enabled easy spatial manipulation when performing a letter-based gesture with the robot arm. We acknowledge that our study design (combining end-effector and gesture) does not allow for investigating effects in isolation. Moreover, we recognize that the number of repetitions in the first and second sessions (n=5), though convenient for the study duration, is rather limited. It remains uncertain if the collected data is sufficient to train data-demanding Deep Learning (DL) classifiers [1]. To that end, we plan to expand our collected data in two directions. First, we plan to investigate different gestures with multiple end-effectors. Second, we consider extending the number of repetitions per condition to be able to investigate the identification accuracy with more extensive data for different classifiers.

6 CONCLUSION

With accuracies between 0.80 and 0.87, we found that the investigated approach – particularly through letter-based gestures – is well-suited as a biometric. In a follow-up study with a subset of participants from the first study, we evaluated our classifiers and reached an accuracy of 0.54. Focusing only on the starting point of interaction, reaching an accuracy score of 0.56 suggests the ability of this approach to generalize to a broader set of gestures.

Our work implies that existing robot arms can be easily extended with user identification functionality, negating the demand for additional hardware. The results also suggest that user identification while interacting with a robot will eventually be feasible through behavioral biometrics. While we explicitly asked participants to perform specific and somewhat complex gestures, this approach can be adapted for implicit identification with simpler, more straightforward gestures (e.g., nudges).

ACKNOWLEDGMENTS

REFERENCES

[1] Amina Adadi. 2021. A survey on data-efficient algorithms in big data era. *Journal of Big Data* 8, 1 (2021), 1–54.
[2] Florian Alt, Stefan Schneegass, Jens Grossklags, Heather Richter Lipford, and Jessica Staddon. 2022. Beyond Passwords—Challenges and Opportunities of

Future Authentication. *IEEE Security & Privacy* 20, 1 (2022), 82–86. https://doi.org/10.1109/MSEC.2021.3127459

[3] Salil Banerjee and D.L. Woodard. 2012. Biometric Authentication and Identification Using Keystroke Dynamics: A Survey. *Journal of Pattern Recognition Research* 7 (01 2012), 116–139. https://doi.org/10.13176/11.427

[4] Wilhelm Bauer, Manfred Bender, Martin Braun, PETER Rally, and Oliver Scholtz. 2016. Lightweight robots in manual assembly–best to start simply. *Frauenhofer-Institut für Arbeitswirtschaft und Organisation IAO, Stuttgart* (2016).

[5] Roger Bemelmans, Gert Jan Gelderblom, Pieter Jonker, and Luc de Witte. 2012. Socially assistive robots in elderly care: a systematic review into effects and effectiveness. *Journal of the American Medical Directors Association* 13, 2 (2012), 114–120. https://doi.org/10.1016/j.jamda.2010.10.002

[6] Aude Billard, Sylvain Calinon, Rüdiger Dillmann, and Stefan Schaal. 2008. Robot Programming by Demonstration. In *Springer Handbook of Robotics.* Springer Berlin Heidelberg, 1371–1394. https://doi.org/10.1007/978-3-540-30301-5_60

[7] Katleen Blanchet, Amel Bouzeghoub, Selma Kchir, and Olivier Lebec. 2020. How to Guide Humans Towards Skills Improvement in Physical Human-Robot Collaboration Using Reinforcement Learning?. In *2020 IEEE International Conference on Systems, Man, and Cybernetics (SMC).* 4281–4287. https://doi.org/10.1109/SMC42975.2020.9283469

[8] Andrea Bonci, Pangcheng David Cen Cheng, Marina Indri, Giacomo Nabissi, and Fiorella Sibona. 2021. Human-robot perception in industrial environments: A survey. *Sensors* 21, 5 (2021), 1571. https://doi.org/10.3390/s21051571

[9] Attaullah Buriro, Bruno Crispo, Filippo Delfrari, and Konrad Wrona. 2016. Hold and sign: A novel behavioral biometrics for smartphone user authentication. In *2016 IEEE security and privacy workshops (SPW).* IEEE, 276–285.

[10] Lipeng Chen, Luis F C Figueredo, and Mehmet R. Dogar. 2018. Planning for Muscular and Peripersonal-Space Comfort During Human-Robot Forceful Collaboration. In *2018 IEEE-RAS 18th International Conference on Humanoid Robots (Humanoids).* 1–8. https://doi.org/10.1109/HUMANOIDS.2018.8624978

[11] Luca De Luisa, Gabriel Emile Hine, Emanuele Maiorana, and Patrizio Campisi. 2021. In-air 3D dynamic signature recognition using haptic devices. In *2021 IEEE International Workshop on Biometrics and Forensics (IWBF).* IEEE, 1–6.

[12] Cengiz Deniz and Gürsel Gökmen. 2022. A New Robotic Application for COVID-19 Specimen Collection Process. *Journal of Robotics and Control (JRC)* 3, 1 (2022), 73–77.

[13] Sandra Drolshagen, Max Pfingsthorn, Pascal Gliesche, and Andreas Hein. 2021. Acceptance of Industrial Collaborative Robots by People With Disabilities in Sheltered Workshops. *Frontiers in Robotics and AI* 7 (2021), 173. https://doi.org/10.3389/frobt.2020.541741

[14] Åsa Fast-Berglund, Filip Palmkvist, Per Nyqvist, Sven Ekered, and Magnus Åkerman. 2016. Evaluating cobots for final assembly. *Procedia CIRP* 44 (2016), 175–180.

[15] Sandra G Hart. 2006. NASA-task load index (NASA-TLX); 20 years later. In *Proceedings of the human factors and ergonomics society annual meeting*, Vol. 50. Sage publications Sage CA: Los Angeles, CA, 904–908.

[16] Sandra G Hart and Lowell E Staveland. 1988. Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In *Advances in psychology.* Vol. 52. Elsevier, 139–183.

[17] Evert Helms, Rolf Dieter Schraft, and M Hagele. 2002. rob@ work: Robot assistant in industrial environments. In *Proceedings. 11th IEEE International Workshop on Robot and Human Interactive Communication.* IEEE, 399–404.

[18] Khoi Hoang Dinh, Ozgur S Oguz, Mariam Elsayed, and Dirk Wollherr. 2019. Adaptation and transfer of robot motion policies for close proximity human-robot interaction. *Frontiers in Robotics and AI* (2019), 69.

[19] Long Huang, Zhen Meng, Zeyu Deng, Chen Wang, Liying Li, and Guodong Zhao. 2021. Extracting Human Behavioral Biometrics from Robot Motions. Association for Computing Machinery, New York, NY, USA, 825–827. https://doi.org/10.1145/3447993.3482860

[20] Long Huang, Zhen Meng, Zeyu Deng, Chen Wang, Liying Li, and Guodong Zhao. 2021. Robot Behavior-Based User Authentication for Motion-Controlled Robotic Systems. In *IEEE INFOCOM 2021-IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS).* IEEE, 1–6.

[21] Satoru Imura and Hiroshi Hosobe. 2018. A Hand Gesture-Based Method for Biometric Authentication. In *Human-Computer Interaction. Theories, Methods, and Human Issues.* Springer International Publishing, 554–566. https://doi.org/10.1007/978-3-319-91238-7_43

[22] Esmeralda Kadena, Lourdes Cecilia Ruiz Salvador, and Zoltan Rajnai. 2022. Behavioral Biometrics for more (dis) trust and security. In *2022 IEEE 16th International Symposium on Applied Computational Intelligence and Informatics (SACI).* IEEE, 000081–000086.

[23] Siddhant Kakade, Bhumeshwar Patle, and Ashish Umbarkar. 2023. Applications of collaborative robots in agile manufacturing: a review. *Robotic Systems and Applications* 3, 1 (jun 2023), 59–83. https://doi.org/10.21595/rsa.2023.23238

[24] Manisha Kamarushi, Stacey Watson, Garreth Tigwell, and Roshan Peiris. 2022. OneButtonPIN: A Single Button Authentication Method for Blind and Low Vision Users to Improve Accessibility and Prevent Eavesdropping. *Proceedings of the ACM on Human-Computer Interaction* 6 (08 2022). https://doi.org/10.1145/3546747

[25] Danica Kragic, Joakim Gustafson, Hakan Karaoguz, Patric Jensfelt, and Robert Krug. 2018. Interactive, Collaborative Robots: Challenges and Opportunities.. In *IJCAI.* 18–25.

[26] Jonathan Liebers, Mark Abdelaziz, Lukas Mecke, Alia Saad, Jonas Auda, Uwe Gruenefeld, Florian Alt, and Stefan Schneegass. 2021. Understanding user identification in virtual reality through behavioral biometrics and the effect of body normalization. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems.* 1–11.

[27] Jonathan Liebers, Sascha Brockel, Uwe Gruenefeld, and Stefan Schneegass. 2022. Identifying Users by Their Hand Tracking Data in Augmented and Virtual Reality. *International Journal of Human–Computer Interaction* (2022). https://doi.org/10.1080/10447318.2022.2120845 https://identifying-users-by-hand-tracking-data.hcigroup.de.

[28] Praveen Kumar Reddy Maddikunta, Quoc-Viet Pham, B Prabadevi, Natarajan Deepa, Kapal Dev, Thippa Reddy Gadekallu, Rukhsana Ruby, and Madhusanka Liyanage. 2022. Industry 5.0: A survey on enabling technologies and potential applications. *Journal of Industrial Information Integration* 26 (2022), 100257.

[29] Eloise Matheson, Riccardo Minto, Emanuele G. G. Zampieri, Maurizio Faccio, and Giulio Rosati. 2019. Human–Robot Collaboration in Manufacturing Applications: A Review. *Robotics* 8, 4 (2019). https://doi.org/10.3390/robotics8040100

[30] Inaki Maurtua, Aitor Ibarguren, Johan Kildal, Loreto Susperregi, and Basilio Sierra. 2017. Human–robot collaboration in industrial applications: Safety, interaction and trust. *International Journal of Advanced Robotic Systems* 14, 4 (2017), 1729881417716010.

[31] Lukas Mecke, Daniel Buschek, Mathias Kiermeier, Sarah Prange, and Florian Alt. 2019. Exploring intentional behaviour modifications for password typing on mobile touchscreen devices. In *Fifteenth Symposium on Usable Privacy and Security (SOUPS 2019).* 303–317.

[32] Jamshed Memon, Maira Sami, Rizwan Ahmed Khan, and Mueen Uddin. 2020. Handwritten optical character recognition (OCR): A comprehensive systematic literature review (SLR). *IEEE Access* 8 (2020), 142642–142668.

[33] George Michalos, Sotiris Makris, Panagiota Tsarouchi, Toni Guasch, Dimitris Kontovrakis, and George Chryssolouris. 2015. Design considerations for safe human-robot collaborative workplaces. *Procedia CIrP* 37 (2015), 248–253.

[34] Robert Miller, Natasha Kholgade Banerjee, and Sean Banerjee. 2022. Temporal effects in motion behavior for virtual reality (VR) biometrics. In *2022 IEEE Conference on Virtual Reality and 3D User Interfaces (VR).* IEEE, 563–572.

[35] Abolfazl Mohebbi. 2020. Human-robot interaction in rehabilitation and assistance: a review. *Current Robotics Reports* 1 (2020), 131–144.

[36] Toshio Morita, Hiroyasu Iwata, and Shigeki Sugano. 2000. Human symbiotic robot design based on division and unification of functional requirements. In *Proceedings 2000 ICRA. Millennium Conference. IEEE International Conference on Robotics and Automation. Symposia Proceedings (Cat. No. 00CH37065),* Vol. 3. IEEE, 2229–2234.

[37] Jonathan Mumm and Bilge Mutlu. 2011. Human-Robot Proxemics: Physical and Psychological Distancing in Human-Robot Interaction. In *Proceedings of the 6th International Conference on Human-Robot Interaction* (Lausanne, Switzerland) *(HRI '11).* Association for Computing Machinery, New York, NY, USA, 331–338. https://doi.org/10.1145/1957656.1957786

[38] Ali Bou Nassif, Ismail Shahin, Imtinan Attili, Mohammad Azzeh, and Khaled Shaalan. 2019. Speech recognition using deep neural networks: A systematic review. *IEEE access* 7 (2019), 19143–19165.

[39] Hyeonjun Park and Donghan Kim. 2020. An open-source anthropomorphic robot hand system: HRI hand. *HardwareX* 7 (2020), e00100.

[40] Max Pascher, Annalies Baumeister, Stefan Schneegass, Barbara Klein, and Jens Gerken. 2021. Recommendations for the Development of a Robotic Drinking and Eating Aid - An Ethnographic Study. In *Human-Computer Interaction – INTERACT 2021* (2021-09-01), Carmelo Ardito, Rosa Lanzilotti, Alessio Malizia, Helen Petrie, Antonio Piccinno, Giuseppe Desolda, and Kori Inkpen (Eds.). Springer, Cham. https://doi.org/10.1007/978-3-030-85623-6_21

[41] Max Pascher, Uwe Gruenefeld, Stefan Schneegass, and Jens Gerken. 2023. How to Communicate Robot Motion Intent: A Scoping Review. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems - CHI '23.* https://doi.org/10.1145/3544548.3580857

[42] Max Pascher, Kirill Kronhardt, Til Franzen, Uwe Gruenefeld, Stefan Schneegass, and Jens Gerken. 2022. My Caregiver the Cobot: Comparing Visualization Techniques to Effectively Communicate Cobot Perception to People with Physical Impairments. *Sensors* 22, 3 (2022), 755.

[43] Luka Peternel, Nikos Tsagarakis, Darwin Caldwell, and Arash Ajoudani. 2016. Adaptation of robot physical behaviour to human fatigue in human-robot co-manipulation. In *2016 IEEE-RAS 16th International Conference on Humanoid Robots (Humanoids).* 489–494. https://doi.org/10.1109/HUMANOIDS.2016.7803320

[44] Ken Pfeuffer, Matthias J. Geiger, Sarah Prange, Lukas Mecke, Daniel Buschek, and Florian Alt. 2019. Behavioural Biometrics in VR: Identifying People from Body Motion and Relations in Virtual Reality. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) *(CHI '19).* Association for Computing Machinery, New York, NY, USA, 1–12. https://doi.org/10.1145/3290605.3300340

[45] Fabio Pini, Francesco Leali, and Matteo Ansaloni. 2015. A systematic approach to the engineering design of a HRC workcell for bio-medical product assembly. In *2015 IEEE 20th Conference on Emerging Technologies & Factory Automation (ETFA)*. IEEE, 1–8.

[46] Jakub Pizoń, Michał Cioch, Łukasz Kański, Eduardo Sánchez-García, et al. 2022. Cobots implementation in the era of Industry 5.0 using modern business and management solutions. (2022).

[47] Henning Pohl, Markus Krause, and Michael Rohs. 2015. One-button recognizer: exploiting button pressing behavior for user differentiation. In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. 403–407.

[48] Alireza Sahami Shirazi, Peyman Moghadam, Hamed Ketabdar, and Albrecht Schmidt. 2012. Assessing the Vulnerability of Magnetic Gestural Authentication to Video-Based Shoulder Surfing Attacks. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Austin, Texas, USA) *(CHI '12)*. Association for Computing Machinery, New York, NY, USA, 2045–2048. https://doi.org/10.1145/2207676.2208352

[49] Maximilian Schrapel, Dennis Grannemann, and Michael Rohs. 2022. Sign H3re: Symbol and X-Mark Writer Identification Using Audio and Motion Data from a Digital Pen. *Mensch und Computer 2022-Tagungsband* (2022).

[50] TaeWon Seo, Youngjae Jeon, Changmin Park, and Jongwon Kim. 2019. Survey on glass and façade-cleaning robots: Climbing mechanisms, cleaning methods, and applications. *International Journal of Precision Engineering and Manufacturing-Green Technology* 6, 2 (2019), 367–376.

[51] Jane Shi, Glenn Jimmerson, Tom Pearson, and Roland Menassa. 2012. Levels of human and robot collaboration for automotive manufacturing. In *Proceedings of the Workshop on Performance Metrics for Intelligent Systems*. 95–100.

[52] Martin Tykal, Alberto Montebelli, and Ville Kyrki. 2016. Incrementally assisted kinesthetic teaching for programming by demonstration. In *2016 11th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE. https://doi.org/10.1109/hri.2016.7451753

[53] Sarah Theres Völkel, Ramona Schödel, Daniel Buschek, Clemens Stachl, Quay Au, Bernd Bischl, Markus Bühner, and Heinrich Hussmann. 2019. Opportunities and challenges of utilizing personality traits for personalization in HCI. *Personalized human–computer interaction* (2019), 31.

[54] Changsheng Wan, Li Wang, and Vir V Phoha. 2018. A survey on gait recognition. *ACM Computing Surveys (CSUR)* 51, 5 (2018), 1–35.

[55] Christian JAM Willemse and Jan BF Van Erp. 2019. Social touch in human–robot interaction: Robot-initiated touches can induce positive responses without extensive prior bonding. *International journal of social robotics* 11, 2 (2019), 285–304.

[56] Sean Ye, Karen Feigh, and Ayanna Howard. 2020. Learning in Motion: Dynamic Interactions for Increased Trust in Human-Robot Interaction Games. In *2020 29th IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)*. 1186–1189. https://doi.org/10.1109/RO-MAN47096.2020.9223437

[57] Yu Zhong and Yunbin Deng. 2015. A survey on keystroke dynamics biometrics: approaches, advances, and evaluations. *Recent Advances in User Authentication Using Keystroke Dynamics Biometrics* 1, 1-22 (2015), 2.

[58] Zuyuan Zhu and Huosheng Hu. 2018. Robot Learning from Demonstration in Robotic Assembly: A Survey. *Robotics* 7, 2 (April 2018), 17. https://doi.org/10.3390/robotics7020017