

## RESEARCH ARTICLE

# NICOL: A Neuro-Inspired Collaborative Semi-Humanoid Robot That Bridges Social Interaction and Reliable Manipulation

MATTHIAS KERZEL<sup>1</sup>, PHILIPP ALLGEUER<sup>1</sup>, ERIK STRAHL<sup>1</sup>, NICOLAS FRICK<sup>1</sup>,  
JAN-GERRIT HABEKOST<sup>1</sup>, MANFRED EPPE<sup>2</sup>, AND STEFAN WERMTER<sup>1</sup>, (Member, IEEE)

<sup>1</sup>Knowledge Technology Group, Department of Informatics, University of Hamburg, 22527 Hamburg, Germany

<sup>2</sup>Institute for Data Science Foundations, Hamburg University of Technology, 21079 Hamburg, Germany

Corresponding author: Philipp Allgeuer (philipp.allgeuer@uni-hamburg.de)

This work was supported in part by the Deutsche Forschungsgemeinschaft (DFG) (CML, MoReSpace, LeCAREbot), in part by the Bundesministerium für Wirtschaft und Klimaschutz (BMWK) (VERIKAS), and in part by the EU Commission (TRAIL, TERAIS).

**ABSTRACT** Robotic platforms that can efficiently collaborate with humans in physical tasks constitute a major goal in robotics. However, many existing robotic platforms are either designed for social interaction or industrial object manipulation tasks. The design of collaborative robots seldom emphasizes both their social interaction and physical collaboration abilities. To bridge this gap, we present the novel semi-humanoid NICOL, the *Neuro-Inspired COLlaborator*. NICOL is a large, newly designed, scaled-up version of its well-evaluated predecessor, the Neuro-Inspired COmpanion (NICO). NICOL adopts NICO's head and facial expression display and extends its manipulation abilities in terms of precision, object size, and workspace size. Our contribution in this paper is twofold—firstly, we introduce the design concept for NICOL, and secondly, we provide an evaluation of NICOL's manipulation abilities by presenting a novel extension for an end-to-end hybrid neuro-genetic visuomotor learning approach adapted to NICOL's more complex kinematics. We show that the approach outperforms the state-of-the-art Inverse Kinematics (IK) solvers KDL, TRACK-IK and BIO-IK. Overall, this article presents for the first time the humanoid robot NICOL, and contributes to the integration of social robotics and neural visuomotor learning for humanoid robots.

**INDEX TERMS** Humanoid robotics, neuro-genetic visuomotor learning, neuro-robotics.

## I. INTRODUCTION

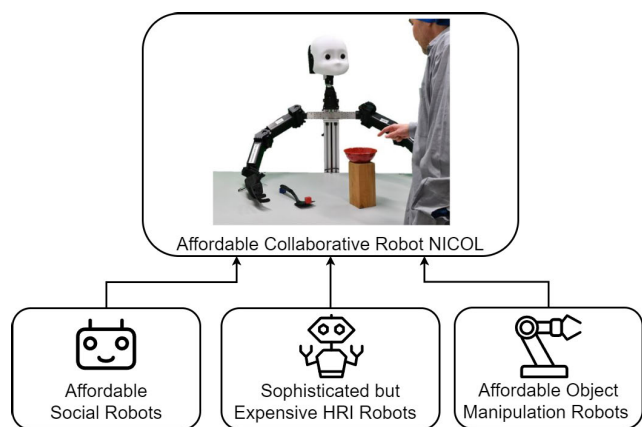
What are the main prerequisites for a collaborative robot to act and work alongside humans? On the one hand, the robot needs to be able to perform precise and reliable object manipulation in a large workspace, but on the other hand, the robot also needs to be able to interact with human coworkers on a social level. This requires auditory communication as well as the understanding and synthesis of both social cues and non-verbal communication like gestures and facial expressions.

Social interaction is one of the crucial factors for intuitive human-robot collaboration and the ability of robots to learn from humans [1], [2], [3]. Gaze, as a non-verbal social cue,

The associate editor coordinating the review of this manuscript and approving it for publication was Yangmin Li.

facilitates shared attention, which can signal future actions or draw attention toward an object or location [4]. Likewise, facial expressions can be used to give feedback or warn about dangers. Speech output also offers an intuitive way to request aid and instructions, or to suggest a course of action. There exist only very few robots, however, that integrate social interaction with reliable manipulation of adult-scale objects while remaining affordable. A robot is desirable that bridges this gap (see Fig. 1). See Section II-A for an overview of related platforms.

The social interaction capabilities of NICOL are based on the well-evaluated design of its predecessor, the humanoid child-sized Neuro-Inspired COmpanion (NICO) [5]. Like NICO, NICOL has an articulated head that can express stylized facial expressions and focus its gaze on interaction partners or relevant elements of the shared workspace.



**FIGURE 1.** NICOL is an affordable robot platform that bridges the gap between social robotics and reliable object manipulation in a large workspace.

While NICO has limited low-precision object manipulation abilities, NICOL can manipulate objects using two Robotis OpenManipulator-P arms with adult-sized five-fingered Seed Robotics RH8D hands attached. It has a workspace similar to an adult person and can handle everyday objects.

In addition to the hardware, NICOL also features a software framework that is directed towards two groups of users—the roboticists that wish to work at a low level with a high degree of control, and the users from different backgrounds, like artificial intelligence, machine learning, cognitive modelling and human-robot interaction, that require a simpler and more abstract interface. Low-level control is delivered via the Robot Operating System (ROS) and provides a rich control interface with great customizability and extensibility. Manipulating and developing within a ROS workspace often requires many non-trivial software cross-dependencies and requires a certain level of expertise to navigate, use and maintain. For users with little background in robotics, a more lightweight system is desired. Furthermore, such users often demand a lightweight simulated version of the robot for performing rapid iterations of parallelized deep reinforcement learning experiments. NICOL meets the demands of both groups—roboticists and AI researchers. Its software offers a full ROS interface, but users can also connect to a front-end Python interface running on top of ROS, which enables them to directly use the robot without having to explicitly deal with ROS. In addition, we provide a simulated virtual version of NICOL.

We introduce NICOL as a novel robot platform; we provide a thorough evaluation of the main advances of NICOL over its predecessor, NICO, its enhanced manipulation abilities and a novel extension to a visuomotor learning approach. In detail, our contributions are:

- We place NICOL in context by examining existing related platforms (Section II). This demonstrates how our robot fills the gap between social robotics and object manipulation in an affordable manner.
- Section III presents the hardware and software concept for NICOL, the *Neuro-Inspired COLlaborator*. Herein,

we also introduce the social interaction capabilities of NICOL, including face and emotion generation capabilities as well as auditory and visual capabilities.

- Section IV presents a thorough evaluation of NICOL's kinematics. We adapt and extend different neural and hybrid neuro-genetic visuomotor approaches initially developed for NICO to the larger NICOL and its more complex kinematics. We show the challenges of neural end-to-end learning for complex kinematics and successfully evaluate a hybrid neuro-genetic approach.
- Section V introduces a novel neuro-genetic visuomotor learning approach that improves the learned grasping accuracy of NICOL to over 99%, and contributes to the growing research on neuro-robotic approaches that are suitable for transfer between platforms with different morphologies [6]. Here, we also present experiments to show that the grasp accuracy of NICOL improves to over 99%, outperforming the state-of-the-art inverse kinematics (IK) solvers KDL, TRACK-IK and BIO-IK.

Overall, this article presents for the first time NICOL, which serves as an affordable research platform that integrates social robotics and humanoid manipulation, and as a testbed for neural visuomotor learning.

## II. RELATED WORK

### A. RELATED ROBOTIC PLATFORMS

NICOL combines the capabilities of dexterous robotic manipulation and social interaction, and so has different types of robots as related platforms.

There are many robotic manipulators for industry and research. Some common platforms are the UR-5 [7] and its related designs from Universal Robots, as well as the KUKA-DLR Lightweight Robot arm [8], and the Franka Research 3 [9]. These robotic arms have at least six degrees of freedom (DoF) to ensure solvable inverse kinematics for general 6-dimensional poses. These manipulators and their various end-effectors are primarily designed to handle objects and lack the social interaction capabilities that enable intuitive learning from, or teaching by, humans.

On the other side of the spectrum are platforms designed for social interaction with no, or limited, manipulation capabilities. For instance, the iCub [10] is available as a stand-alone 3-DoF head that can display facial expressions and perform gaze shifts. The Furhat [11] robot head is another example and can project animated or recorded faces for interaction purposes.

Infant-sized or toy-sized humanoids form another type of social platform. Their small size is advantageous for research, they have an affordable price, and they are inherently safe due to low motor speeds and weight. Well-known examples are the Softbank Robotics NAO (formerly Aldebaran) and the ROBOTIS OP3 (descendant of the DARwIn-OP) [12]. While infant-sized platforms often have humanoid manipulators, they are too small to manipulate adult-sized items or reach them in domestic-scale environments.

**TABLE 1. Robotic platforms for social interaction and object manipulation. Price ranges are estimated; a low price range is below 10,000 EUR, a medium price range is below 100,000 EUR and a high price range is above this. Few platforms are designed for both social interaction and physical tasks while still being affordable. NICOL fills the gap of a robotic platform with both an upper humanoid body and humanoid head, capable of facial expressions, gestures and object manipulation in an adult-sized workspace at an affordable price tag.**

Robot	Type	Humanoid Body	Humanoid Head	Facial Expressions	Gestures	Grasping	Workspace	Price Range
UR-5	robot arm	✗	✗	✗	✗	✓	adult-sized	mid
KUKA-DLR Arm	robot arm	✗	✗	✗	✗	✓	adult-sized	mid
Franka Research 3	robot arm	✗	✗	✗	✗	✓	adult-sized	low
Furhat	robot head	✗	✓	✓	✗	✗	none	mid
iCub head	robot head	✗	✓	✓	✗	✗	none	mid
ROBOTIS OP3	infant-sized robot	✓	✗	✗	✓	✗	tiny	low
Nao	infant-sized robot	✓	✓	✗	✓	✓	tiny	low
iCub	child-sized robot	✓	✓	✓	✓	✓	child-sized	high
NimbRo-OP	child-sized robot	✓	✓	✗	✗	✗	child-sized	mid
Pepper	child-sized robot	(✓) upper body	✓	✗	✓	✗	child-sized	mid
Poppy	child-sized robot	✓	✓	✓	✓	✓	child-sized	low
NICO (ours)	child-sized robot	✓	✓	✓	✓	✓	child-sized	low
Talos	adult-sized robot	✓	✓	✗	✓	✓	adult-sized	high
PR2	adult-sized robot	✗	✗	✗	✗	✓	adult-sized	high
Sawyer	adult-sized robot	✗	✗	✓	✗	✓	adult-sized	mid
ARMAR-6	adult-sized robot	(✓) upper body	✗	✗	✓	✓	adult-sized	high
TOMM	adult-sized robot	(✓) upper body	✗	✗	✗	✓	adult-sized	high
Fetch	adult-sized robot	✗	✗	✗	✗	✓	adult-sized	high
R1	adult-sized robot	(✓) upper body	✗	✓	✓	✓	adult-sized	high
Sophia	adult-sized robot	(✓) upper body	✓	✓	✓	✗	adult-sized	high
NICOL (ours)	adult-sized robot	(✓) upper body	✓	✓	✓	✓	adult-sized	mid

The next larger category of humanoids is child-sized. A popular platform for developmental research is the iCub [10]. Its size resembles a child of about 90 cm, it is well-actuated with 53 DoF, features human-like hands and gaze shifts, and can display stylized facial expressions. NICOL is significantly more affordable and simple than the iCub, and is thereby less challenging to modify and/or customize for particular needs. The NimbRo-OP family of robots [13], [14], [15] are designed for the RoboCup soccer league, and focus on bipedal walking with arms for balance as well as getting up, but not on manipulation. As such, they do not even have hands that can be actuated. Softbank’s Pepper robot [16] features a humanoid torso on top of an omnidirectional wheeled platform. It has 20 DoF in the upper body, but its arms and hands are designed for gesturing, not grasping. Poppy is an open-source 3D-printed robot designed for education, artists, and scientists [17]. It is completely customizable and comes with a display to show facial expressions and other information. Poppy can perform bipedal locomotion, but in its primarily advertised configuration, it features only very limited grasping functionalities. The concept of NICO is quite similar to that of Poppy. NICO is an open platform for researching neuro-robotic models for human-robot interaction, as well as visuomotor learning [5]. Its head adapts the open iCub design, and, in contrast to Poppy, integrated LED arrays display stylized facial expressions. Furthermore, NICO has quite sophisticated grasping abilities to manipulate small objects, provided by two 6-DoF arms with fully functional child-sized anthropomorphic hands produced by Seed Robotics.

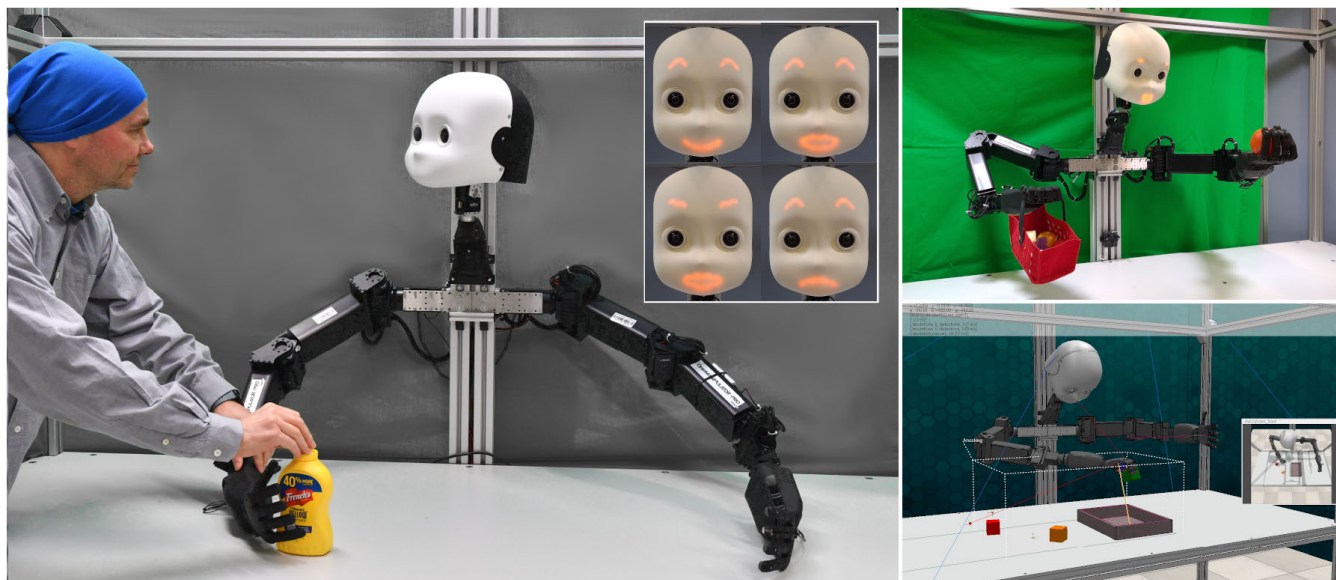
Adult-sized humanoids with manipulation abilities can generally handle real-life objects. However, such platforms are often expensive and difficult to maintain—like PAL’s Talos [18], which was introduced at a price of about

1 million euros—or their design is too non-humanoid for social interaction, like the Atlas initially developed by Boston Dynamics, and the PR2 from Willow Garage. Some adult-sized platforms lack object manipulation abilities entirely, like the Hanson Robotics Sophia robot. The ARMAR-6 from the KIT [19] is designed for complex collaborative tasks in industrial environments, but it neither features a human-like face nor can it display facial expressions. Likewise, the TOMM [20] and the one-armed Fetch robot [21] focus on manipulation and wheeled mobility, but not on social cues. The R1 from the IIT [22] is a wheeled platform with grippers that can display an animated face on a screen. A related design is the Rethink Robotics Sawyer, the successor of the well-known humanoid Baxter robot. It features a tablet for displaying animated eyes, but it only has a single end effector and is not humanoid.

Table 1 summarizes our analysis of robotic platforms. It is evident that there is a gap in available platforms that are designed for both social interaction and adult-sized object manipulation tasks. The few platforms that fulfill these two criteria are prohibitively expensive. We propose the open NICOL design to address this gap in the state of the art. In this article, we show that the NICOL is capable of social interaction tasks by extrapolating from previous work on the NICO, its “smaller brother”. NICO and NICOL share the design of their head and facial expression mechanism(s), but importantly, the latter possesses a significantly increased arm strength and workspace size, allowing for more real object scenarios.

### III. SEMI-HUMANOID ROBOTIC PLATFORM NICOL

NICOL is designed for research on robots that learn from and collaborate with humans. It can serve as a platform for integrating neuro-robotic visuomotor models combined with



**FIGURE 2.** NICOL semi-humanoid platform. Left: NICOL jointly works on a grasping task with an experimenter (in the upper-right corner thereof, different facial expressions of NICOL are shown that can give feedback during the task). Right: NICOL object manipulation and expression in real life (top) and simulation (bottom).

affective interaction, joint attention, and shared perception. NICOL, depicted in Fig. 2, consists of three main components: two manipulators with anthropomorphic hands, and the head with a facial expression display. NICOL adapts the head design of the child-sized humanoid NICO and combines it with two Robotis manipulators with five-fingered hands. NICOL can naturally collaborate with human partners on physical tasks, thanks to its ability to manipulate adult-size domestic objects and to use non-verbal cues like gaze or facial expression.

#### A. SOCIAL INTERACTION CAPABILITIES OF NICOL

The interaction capabilities of NICOL are based on the design of NICO. Here, we summarize the main findings and studies. The seven facial expressions of NICO, which are now shared by NICOL, are evaluated by Churamani et al. [23]. In a study with twenty participants from eleven different countries, five expressions (neutral, happiness, sadness, surprise and anger) are recognized by the participants with an accuracy of over 75%. The positive effect of the emotion display on the robot's subjective user rating is verified with a Godspeed questionnaire [24]. The freely programmable LED arrays are also used in learning emotion expression via reinforcement learning [25]. Ng et al. [26] use facial expressions in combination with neurocognitive models for social cue recognition and behavior control, and report that a natural language dialogue system benefits from the robot directly looking at the face of its interaction partner. Together with a more personalized conversation, this behavior made participants perceive NICO as more intelligent and likable. Beik-Mohammadi et al. [27] show that using social gestures and more socially engaging dialogue enhanced the robot's

perceived likeability and animacy. Finally, Kerzel et al. [3] show that participants guided by NICO in a visuomotor learning scenario via verbal requests and facial expressions, as compared to instructions by a human experimenter, rated the human-robot interaction as more positive and engaging. This increased engagement can improve the outcome of the learning scenario.

In summary, the studies strongly suggest that the design of NICOL's head, which is adapted from NICO, can effectively create social cues in terms of facial expressions and gaze. They also show that these social cues positively affect human-robot interactions.

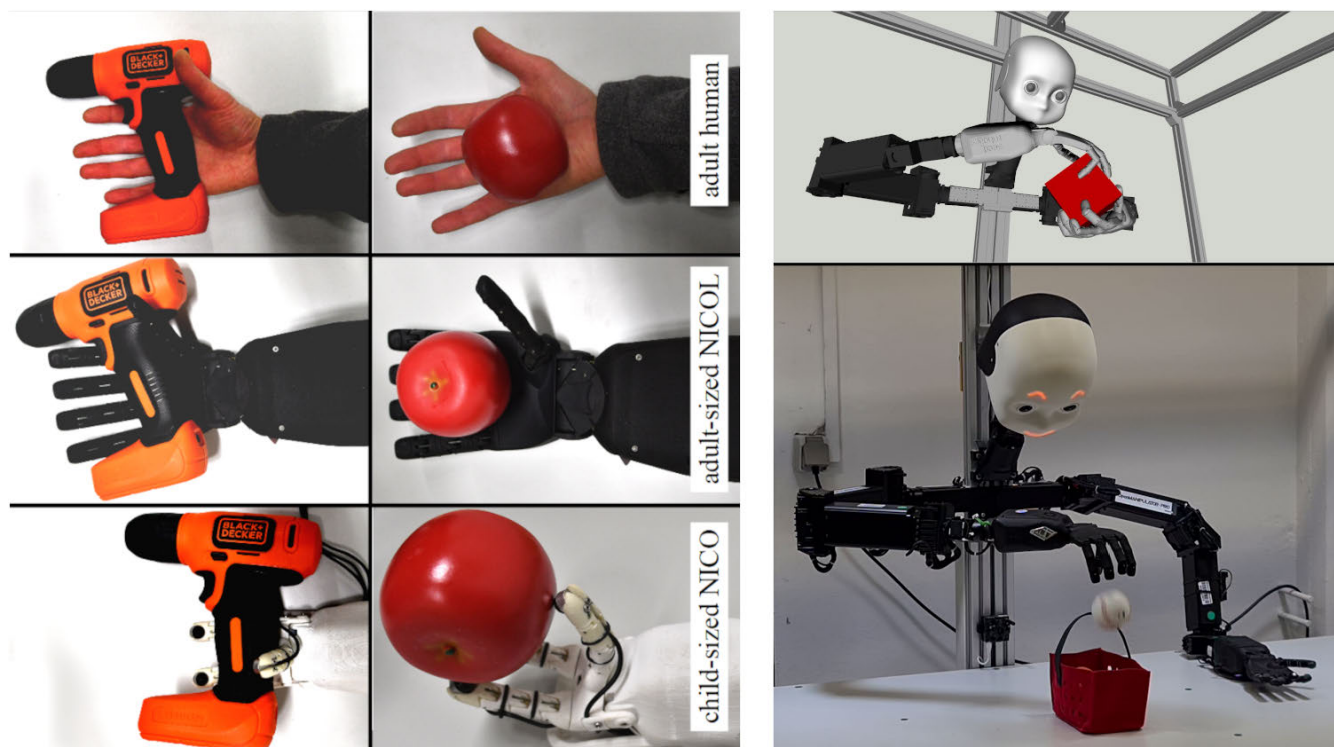
#### B. ARMS AND HANDS

The arms of NICOL consist of two Robotis OpenManipulator-P<sup>1</sup> arms with 6 DoF and a payload of 3 kg. As end-effectors, two SeedRobotics RH8D adult-sized robot hands<sup>2</sup> with a manipulation payload of 750 g are used. All five fingers in the hand are tendon-operated. Each three-segment finger is operated by a single tendon. Each hand has eight actuated DoF—three in the wrist for rotation,<sup>3</sup> flexion, and abduction, two DoF in the thumb for abduction and flexion, and one DoF each for index finger flexion, middle finger flexion and combined flexion of the ring and little finger. The tendon mechanism allows the hand to coil around objects without further fine control. The arms of NICOL can reach up to 100 cm and therefore have a workspace slightly larger than that of an adult sitting at a table.

<sup>1</sup><https://www.robotis.us/openmanipulator-p>

<sup>2</sup><https://www.seedrobotics.com/rh8d-adult-robot-hand>

<sup>3</sup>Note that the wrist rotation is redundant and therefore not used.



**FIGURE 3.** Left: Size comparison between an adult human’s hand, NICOL, and NICO. Upscaling the hands of NICO allows NICOL to manipulate everyday objects and tools. Right: NICOL in a virtual Gazebo environment during a dexterous manipulation task and dropping a baseball into a basket.

### C. HEAD AND AUDIOVISUAL SENSING

The design of the shell of NICO’s head is based on a modified version of the open-source iCub design [10]. The design balances human features with enough abstraction to avoid the uncanny valley effect [28]. Stylized facial expressions can be displayed by NICOL using three LED arrays that are placed behind the eyes (two  $8 \times 8$  LED arrays) and mouth (one  $16 \times 8$  array). Due to the head’s material properties, the individual LEDs form lines shining through the head’s material, as shown in Fig. 2. An internal speaker is used to facilitate spoken communication. The head is articulated with two DoFs for pitch and yaw movements. Two See3CAM CU135 cameras with a  $4096 \times 2160$  (4K) resolution are mounted in the eye sockets of NICOL. Their fisheye lens has a field of view of 202 degrees. Two Soundman OKM II binaural microphones are placed at the sides of NICOL’s head; due to the absence of head-internal fans or mechanics, the robot’s ego noise is very low.

### D. TABLE ENVIRONMENT AND SAFETY MEASURES

The table environment measures  $100 \times 200$  cm and is located at a height of 74 cm. The head and arms of the NICOL are mounted to a vertical support at the rear center of the table. Scaffolding at the corners of the table creates a visible delimitation for the workspace of the robot. A human interaction or collaboration partner is thereby implicitly made aware when actively reaching into the robot’s workspace. For additional safety, two emergency shutdown buttons are

integrated within the operator’s workspace. The scaffolding can further be used to mount external sensors or devices, like cameras or light sources.

### E. SOFTWARE AND SIMULATION ENVIRONMENT

The NICOL API is based on the Robot Operating System (ROS) middleware. All functionalities are provided through a modified version of the hardware controller delivered with the Robotis arms, extended to support the Seed Robotics hands, additional sensors, as well as drivers for custom hardware (facial expressions). The API integrates the MoveIt planning framework and provides a Python-based client. For prototyping and development, as well as for simulation of real-world scenarios, NICOL is realized in both the Gazebo simulation environment (see Fig. 3) and CoppeliaSim (see Fig. 2). The robot description is based on a URDF model, making it possible to import NICOL in many other simulators and environments.<sup>4</sup>

## IV. VISUOMOTOR LEARNING AND STATE-OF-THE-ART IK SOLVERS

To evaluate NICOL as a research platform for neuro-robotic visuomotor learning, we first transfer an end-to-end neural visuomotor learning approach from NICO, a smaller humanoid robot, to NICOL. Based on the results,

<sup>4</sup>Software and CAD/URDF files, including everything needed to run the CoppeliaSim simulation, are available at <https://www.inf.uni-hamburg.de/en/inst/ab/wtm/research/neurobotics/nicol.html>

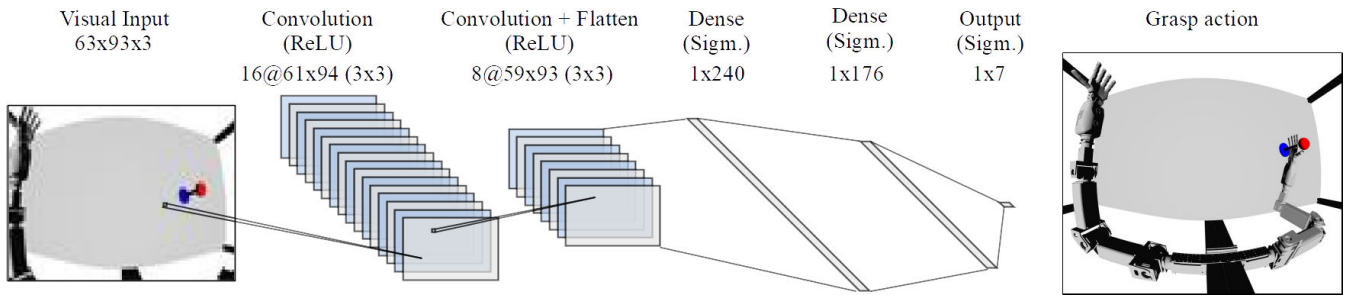


FIGURE 4. Neural architecture for end-to-end visuomotor learning.

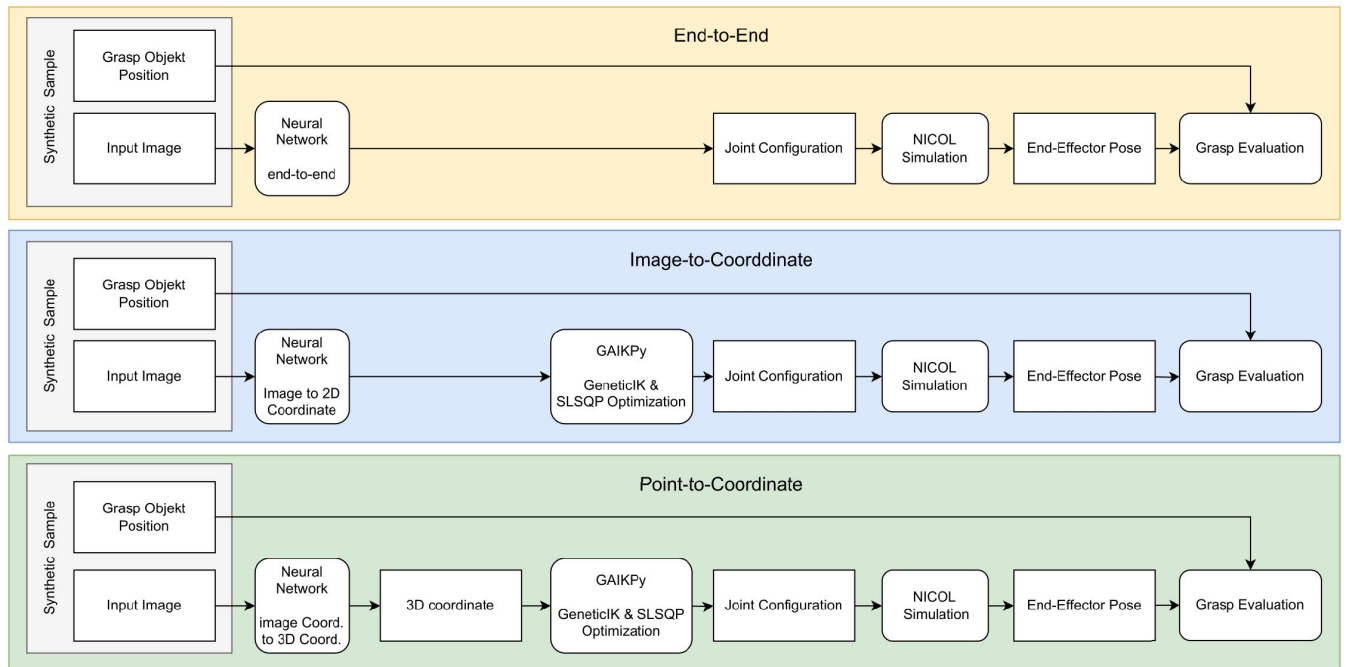


FIGURE 5. Three architectures for visuomotor learning.

we propose two variations of a modular neuro-hybrid approach. In these modular approaches, we use a neural network for image processing and different IK solvers for the kinematics. We evaluate the baseline performance of state-of-the-art IK solvers that are established in the robotics community and suggest a novel hybrid neuro-genetic approach that can handle the added kinematic complexity of a humanoid.

**A. STATE-OF-THE-ART IK APPROACHES**

Solving IK is a well-known and studied problem in the robotics community. The *MoveIt planning framework* [29], as part of the ROS middleware, offers the required flexibility to utilize different IK solvers for an experimental setup. It enables a plug-in-based configuration of inverse kinematics solvers as well as high-level motion planners. Mainly third-party kinematics libraries, contributed by foundations or private sector organizations are available, and these have become state-of-the-art due to the popularity of the ROS

framework. The inverse kinematics and planning functionalities are provided in the form of ROS services and offer a variety of parameterized kinematics calculations.

The *Kinematics and Dynamics Library (KDL)* by Orocos [30] is de facto the most widely adapted inverse kinematics solution in the ROS ecosystem. It supports kinematic chains with a minimum of 6 DoF. KDL utilizes the well-known pseudo-inverse Jacobian method to determine suitable joint configurations. As the Jacobian matrix holds the partial derivatives between joint space and Cartesian space, the difference between the so-called seed state (typically the initial joint configuration) and the target state can be calculated in Cartesian space. Forward kinematics are used on the seed state to determine the current end-effector pose. A suitable joint configuration for achieving the target Cartesian end effector pose can then be obtained by calculating the error between the initial pose and the target pose, and multiplying it with the inverted Jacobian. The algorithm makes use of the Newton-Raphson method and iteratively

minimizes the error between the previous and next target poses.

With the intention of improving the coverage of the robot workspace, TracLabs identified key issues in the KDL framework and proposed *TRAC-IK* in 2015 [31]. Remarkable about the approach is that two IK solvers are executed in parallel. One is a modified reimplementation of KDL called KDL-RR. The method additionally checks for local optima during the iterative improvement step and moves the next seed far enough away from them. The second solver, SQP-SS, formulates the IK problem as a non-linear sequential quadratic programming optimization problem. *TRAC-IK* returns the solution of the solver that first terminates with a valid joint configuration. In contrast to the pseudo-inverse Jacobian, SQP-SS is particularly capable of directly handling constraints, such as the joint limits, during the optimization step. This ability is particularly relevant for humanoid robots like NICOL, whose joints are constrained to mimic the human range of motion.

Starke et al. introduced the *Bio-IK* method [32], [33]. The algorithm combines multiple bio-inspired optimization methods, most importantly evolutionary and particle swarm optimization, in order to solve the inverse kinematics problem. In difference to KDL and *TRAC-IK*, it does not rely on any Jacobian mathematics. The evolutionary algorithm is fundamental to the approach. Every individual in the population corresponds to a joint space robot pose. Momentums are assigned to every individual by hybrid particle swarm optimization, and these momentums are continuously updated during runtime. Besides selection, mutation and recombination, also elitism is used to prevent the deterioration of already-found solutions. Local search is executed on the elites, and simulates mutations on single genes. The classical MoveIt interface only allows setting the pose or position goals. *Bio-IK* additionally offers extended functionality in a separate service [34], where multiple custom goal types are available, e.g. linear end-effector trajectories or minimal joint displacement configurations. The various custom goal types can be combined into a single request, as each custom goal type is treated as a weighted partial cost function by the algorithm. The service also offers an approximate mode, leading to lower precision in the IK solutions but enabling higher coverage of the workspace.

### B. END-TO-END NEUROROBOTIC VISUOMOTOR LEARNING

The supervised neural end-to-end visuomotor approach, first introduced in [35], learns to map a single object's visual input in the robot's workspace to joint configurations for reaching for the object. The approach was evaluated with a grasp success rate of approximately 85% on NICO with 5 DoF in a workspace of size  $30 \times 40$  cm and a training set of 400 samples. In this paper, we evaluate if the larger workspace and more complex NICOL kinematics influence the learning outcome. The neural architecture is shown in Figure 4. In Section V-B we report experimental results.

### C. HYBRID NEURO-GENETIC VISUOMOTOR LEARNING

Genetic algorithms are based on evolutionary selection, recombination and mutation processes of a population of individuals, modelled by their chromosomes [36]. Each chromosome encodes a potential solution to a task. Following Kerzel et al. [37], the chromosomes encode a joint configuration to reach a given pose. During each iteration, these chromosomes, expressed as individuals, are ranked according to their fitness, i.e., the position and orientation error compared to the goal pose. Successful individuals are copied into the next generation, with possible random mutations to their chromosome-encoded joint configuration. Niching is used to preserve diverse chromosomes; on every CPU of the computer system, a different evolution is implemented, thus fully utilizing multiple cores with minimal overhead.

While genetic algorithms are adept at escaping local minima, they often have issues further optimizing found solutions. To overcome this limitation, we hybridize the genetic algorithm with gradient-based Sequential Least Squares Programming (SLSQP). This computationally expensive procedure is only applied to the n-best individuals of the population. While preliminary experiments have indicated that SLSQP is prone to be attracted by local minima, the initialization with solutions found by the genetic algorithm can overcome this shortcoming. Optimized hyperparameters for the genetic algorithm and SLSQP were adapted from [37], where they yield good results for NICO.

While the hybrid genetic algorithm provides IK solutions, a neural component is utilized to determine the position of the grasp object from visual input. The architecture is shown in Fig. 5, and experiments are reported in Section V-B.

### V. IK AND VISUOMOTOR LEARNING EXPERIMENTS

We first curate a new dataset (Section V-A) for visuomotor learning on the NICOL platform. A first analysis of the dataset allows us to get an insight into NICOL's workspace. In Section V-B, we evaluate how well our neural end-to-end and a hybrid neuro-genetic visuomotor learning approach can be transferred from the child-sized humanoid NICO to the adult-sized NICOL. We evaluate how the significantly increased workspace and the more complex manipulator affect the learning outcome. In Sections V-C and V-D, we present a modular neuro-hybrid approach in which we evaluate state-of-the-art IK solvers and a neuro-generic approach based on GAIKPy.

#### A. DATASET FOR VISUOMOTOR LEARNING ANALYSIS OF NICOL'S WORKSPACE

We collect three datasets,<sup>5</sup> each consisting of 10,000 samples in a virtual environment. The datasets differ in the number of active joints of NICOL, ranging from six to eight. As the joints of NICOL are constrained to mimic the human range of motion, we hypothesize that the additional degrees

<sup>5</sup>The datasets are available at <https://www.inf.uni-hamburg.de/en/inst/ab/wtm/research/corpora.html>

of freedom beyond six will increase the number of reachable grasp poses. Each sample contains the following: 1) an image from the egocentric perspective of NICOL with a grasp object placed onto the table on the right side of the workspace in a  $100 \times 100$  cm area; the images are cropped and resized to  $63 \times 96$  pixels. 2) The x- and y-coordinates of the grasp object on the table. 3) The configuration of NICOL's 8 arm joints that result in a grasp pose. The grasp pose has the hand pointing forward with the palm touching the grasp object at the middle of its height, as shown in Fig. 4 (right side). We use a genetic algorithm *GAIKPy* introduced in [37] for computing the joint configuration for grasping the object that is placed at a random position within the robot's workspace. If no suitable joint configuration can be found with the genetic algorithm, the sample is rejected, and another random object position is generated.

Fig. 6 shows the distribution of all successful grasp samples in NICOL's workspace for the three experimental conditions. Due to the rigid mechanical constraints of our defined grasping pose, the joint configuration and self-collision avoidance, the area within the defined workspace for which grasp poses can be found increases with the number of active joints. Only with eight active joints, good coverage of the workspace can be achieved. Therefore, we focus our analysis on this dataset.

### B. END-TO-END NEUROROBOTIC VISUOMOTOR LEARNING EXPERIMENTS

First, we evaluate a neural end-to-end learning approach using the neural architecture adopted from [35]. The input to the neural network is a  $96 \times 64$  pixel image from NICOL's perspective of its workspace with a single object placed in it. Depending on the dataset, the output of the network is the seven or eight joint values of NICOL's right manipulator. The hyperparameters were optimized using Hyperopt [38]. The model was trained with an Adam optimizer and a batch size of 35, and the MSE over the joint configuration as a loss function on a random 90-10 split of the entire dataset. Hyperparameter ranges and results are shown in Table 2, and the resulting architecture is shown in Fig. 4.

**TABLE 2. Ranges and results of hyperparameter optimization of the end-to-end visuomotor architecture for 8 DoF.**

Parameter	Parameter range	Result 8 DoF
# filters conv <sub>1</sub>	8-32 (step size 8)	32
# filters conv <sub>2</sub>	8-32 (step size 8)	8
kernel size conv <sub>1</sub>	3-5 (step size 1)	3
kernel size conv <sub>2</sub>	3-5 (step size 1)	5
# dense <sub>1</sub> neurons	16-1024 (step size 16)	256
# dense <sub>2</sub> neurons	16-256 (step size 16)	160
normalisation	[0..1] or [-1..1]	[0..1]
image size	48x32, 96x64, 192x128	96x64
learning rate	[0.01, 0.001, 0.0001, 0.00001]	0.0001

We use 10-fold cross-validation to obtain our results. For evaluation, we denormalize the joint output values and calculate the robotic hand's pose using the known forward kinematics of NICOL. The mean position error

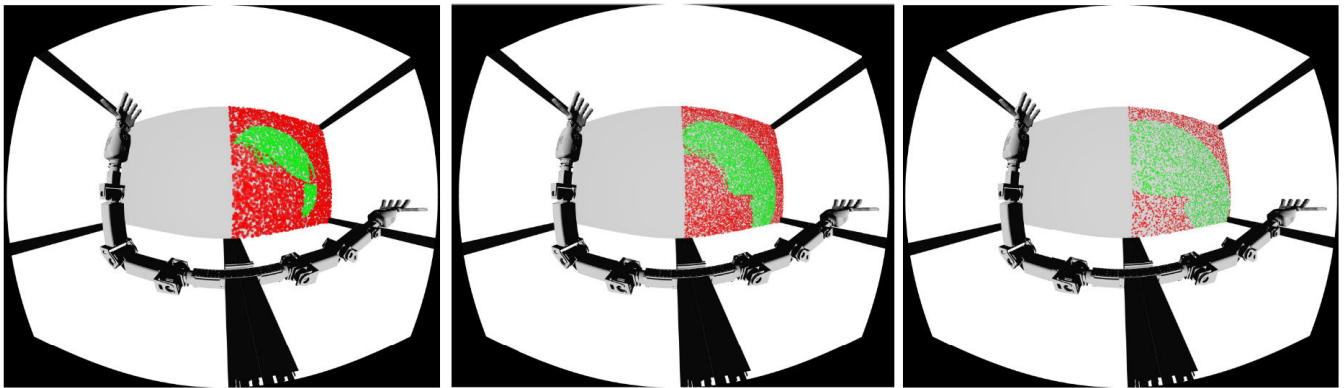
is  $0.42 \pm 0.21$  meters and the mean orientation error is  $7.72 \pm 5.88$  degrees. Based on previous work [35], [37], we count a grasp as successful if the position error of the resulting pose is  $< 10$  mm and the sum of orientation errors is  $< 20$  degrees. These limits are regularly exceeded. Table 3 summarizes the results with a single-digit grasp accuracy. We conclude that the end-to-end approach cannot be transferred directly to NICOL. We attribute this result to several factors: The workspace of NICOL with  $100 \times 100$  cm is larger than the workspace of NICO with  $30 \times 40$  cm; more importantly, the arms of NICOL afford more complex motions and more diverse joint configurations during grasping.

### C. MODULAR APPROACH WITH CLASSICAL IMAGE PROCESSING, NEURAL COORDINATE TRANSFORMATION AND IK SOLVERS

Next, we analyze if the low accuracy of neural end-to-end visuomotor learning can be attributed to issues in visually locating the grasp object in the image, transforming the image coordinates to world coordinates or the computation of inverse kinematics. In previous work, we address this issue using a modular, hybrid neuro-genetic approach [37]. We adopt this approach to the NICOL and use three modules as shown in Fig. 5: Classical image processing to locate the object in the image, a neural-network to transform the image coordinates into world coordinates, and existing state-of-the-art and novel generic IK solvers. To extract the grasp object's position from the non-downsampled image ( $4208 \times 3120$ ) using classical image processing, we detect the top and the bottom of the grasp object with a standard color-based detector, resulting in the x, y position and the radius of the base and top parts of our grasp object. These 6 input parameters are fed into an MLP, which outputs the x- and y-coordinates of the object on the table. The network consists of one dense layer with 180 neurons and ReLU activation; the architecture is based on [37] and was optimized with Hyperopt [38]. Table 4 shows the optimization ranges and results.

We use 10-fold cross-validation for all reported results. First, we evaluate the error of the transformation. As shown in the first line in Table 5, the resulting mean position error is  $0.0039 \pm 0.0023$  m, and the median position error is 3 mm. Applying the above-established criteria for a successful grasp (obviously without considering orientation) resulted in a grasp accuracy of 98.47%. Next, we evaluate the complete grasp architecture by using IK solvers to compute the robot's joint configuration based on the predicted object coordinates. Table 5 summarizes the results for the state-of-the-art IK solvers KDL, TRACK-IK and BIO-IK as well as for our novel genetic IK solver *GAIKPy*. Again, we apply the criteria for a successful grasp (position error is  $< 10$  mm, and the sum of orientation errors is  $< 20$  degrees). We achieved accuracies of 90.52%, 90.29% and 89.60% for KDL, TRAC-IK and BIO-IK respectively. In both conditions, *GAIKPy* shows the highest accuracy with 98.45% and 92.85%.





**FIGURE 6.** Left: Placement of the 10,000 collected samples (green) versus the rejected samples (red) for 6 (left), 7 (middle) and 8 (right) DoF. The images show the continuity of the IK map, i.e., to what extent a smooth IK trajectory ends up as a smooth joint trajectory if each instantaneous pose is individually passed through the IK algorithm.

**TABLE 3.** Results for the neural end-to-end approach for 8 DoF. Results indicate that the neural end-to-end approach cannot be transferred to the more complex kinematics of NICOL.

Experiment	DoF	Mean pos. error	Median pos. Error	Mean orient. error	Median orient. error	Grasp accuracy
End-to-end	8	$0.4161 \pm 0.2085$ m	0.3908 m	$7.72 \pm 5.88$ deg.	6.01 deg.	0.01 %

**TABLE 4.** Ranges and results of hyperparameter optimization of the neural image-to-coordinate transformation.

Parameter	Parameter range	Result
# dense neurons	10-200 (step size 10)	180
normalisation	[0..1] or [-1..1]	[0..1]

#### D. NOVEL NEURAL IMAGE-TO-COORDINATE TRANSFORMATION AND GENETIC ALGORITHM

To evaluate the limits of neural visuomotor learning, we modify the end-to-end architecture presented above to regress directly the x- and y-coordinates of the object on the table instead of the joint configurations. We then use state-of-the-art IK solvers and GAIKPy to compute a joint configuration that reaches for the output coordinate with a given hand orientation.

We use 10-fold cross-validation to obtain the results in Table 6. Applying the criteria for a successful grasp, 96.51%, 96.2% and 95.71% for KDL, TRAC-IK and BIO-IK. GAIKPy shows the highest accuracy with 99.17%.

#### E. DISCUSSION

In Section V-A, we show that 8 DoFs are needed to cover sufficiently large parts of NICOL's intended workspace. In contrast, non-humanoid industrial robot arms often only need 6 DoF to cover their workspace. We attribute this finding to the NICOL design mimicking human motion, resulting in a more humanoid distribution of DoF along the arm and, more importantly, severe constraints on its joints. Next, we show that the increased kinematic complexity of NICOL cannot be handled by an end-to-end neural visuomotor learning approach in Section V-B. However, we successfully apply

a hybrid modular approach using image processing and a neural coordinate transformation with an IK solver in Section V-C. We also demonstrate that NICOL integrates well with the state-of-the-art IK solvers KDL, TRACK-IK and BIO-IK. However, the best results were achieved by using our genetic algorithm-based GAIKPy IK solver. We attribute this finding to the combination of more than 6 DoF in conjunction with joint constraints, posing a challenge for traditional state-of-the-art IK solvers, as shown in Tables 5 and 6. Finally, in Section V-D, we positively evaluated a novel hybrid approach that uses a convolutional neural network to extract target grasp positions from an image in conjunction with an IK solver. Again, our GAIKPy approach yields the highest grasp accuracy with a success rate of over 99%.

This added accuracy comes at the cost of longer processing times. We give each IK solver 1 second to compute a solution. While GAIKPy fully utilizes this time, the state-of-the-art solvers show a different behavior: In cases, where they can find a good solution, this often happens much faster, however, they don't utilize the time budget to find suitable solutions for difficult poses.

We argue that in human-robot collaboration, these slower IK solutions are less problematic. For a humanoid robot, collaborating with a human, we face different challenges compared to an industrial robotic arm. For human-robot collaboration, safe and more importantly, predictable motions are essential. For safety reasons, we need to limit the technical possible maximum joint speed of the humanoid anyway. The slower but more accurate hybrid neuro-genetic approach yields better accuracy scenarios, while still fulfilling the time constraints for collaborative human-robot scenarios.

**TABLE 5. Results for the modular approach with Classical Image Processing, Neural Coordinate Transformation and IK solvers for 8 DoF. GAIKPy achieves the highest grasp accuracy of all evaluated IK solvers.**

Experiment	DoF	Mean pos. error	Median pos. Error	Mean orient. error	Median orient. error	Grasp accuracy
transformation	-	0.0039 ± 0.0023	0.00343	-	-	98.47
KDL	8	0.0229 ± 0.1031	0.00444	0.519 ± 2.924	1.817e-05	90.52
TRAC-IK	8	0.0244 ± 0.1068	0.00445	0.564 ± 3.042	1.630e-05	90.29
BIO IK	8	0.0280 ± 0.1144	0.00448	0.686 ± 3.343	1.987e-04	89.60
GAIKPy	8	0.0067 ± 0.0302	0.00436	0.337 ± 0.687	0.1186	<b>92.85</b>

**TABLE 6. Results for our novel hybrid approach combining neural image-to-coordinate transformation and genetic IK with GAIKPy for 8 DoF. Again, GAIKPy achieves the highest grasp accuracy.**

Experiment	DoF	Mean pos. error	Median pos. Error	Mean orient. error	Median orient. error	Grasp accuracy
transformation	-	0.0018 ± 0.0012	0.00158	-	-	99.98
KDL	8	0.0191 ± 0.1034	0.00125	0.487 ± 2.834	1.837e-05	96.51
TRACK-IK	8	0.0209 ± 0.1078	0.00125	0.540 ± 2.979	1.627e-05	96.20
BIO-IK	8	0.0233 ± 0.1130	0.00126	0.625 ± 3.196	2.035e-04	95.71
GAIKPy	8	0.0019 ± 0.0102	0.00122	0.465 ± 1.151	0.1841	99.17

## VI. CONCLUSION AND OUTLOOK

We present NICOL, a novel semi-humanoid robotic platform designed for research in social robotics and physical human-robot collaboration. NICOL, combining the well-received social interaction features of the NICO platform [5] with adult-sized manipulators, fills a gap in the current state of the art. It offers an affordable and open platform for research involving robots collaborating with and learning from humans in scenarios demanding advanced object manipulation and interaction abilities.

Our evaluation of NICOL primarily focuses on its inverse kinematics capability in the context of a reach-for-grasp task, which is an essential skill for humanoid robots. We positively evaluate our main design goal for NICOL as a successor of the child-sized NICO, and to create a robotic platform that has a workspace and can handle objects comparable to an adult human, thus being able to be used in physical collaboration scenarios. We also address the issue of standard IK solvers having problems with kinematic chains that have more than 6 DoF, while at the same time having strict constraints on the individual joint limits to mimic the human range of motion. We evaluate neural end-to-end learning approaches for visuomotor learning and compare state-of-the-art IK solvers against a novel hybrid neuro-genetic IK approach. We show that neural end-to-end learning is challenging due to the more complex kinematics of NICOL, and overcome this challenge with a neuro-genetic approach. Furthermore, we demonstrate that our novel hybrid neuro-genetic approach outperforms classical IK solvers in terms of accuracy by taking advantage of the relaxed time constraints in human-robot collaboration scenarios.

In future work, we will thoroughly evaluate NICOL's abilities in different collaborative real-world tasks where it will learn from and assist human participants. One such task is assembly, where NICOL will be asked to pick up and hand over different components to a human participant. Besides the motoric challenge, such tasks required shared attention on objects, social cues regarding the handover and verbal communication to coordinate efforts. A second research scenario is learning from a human through observations

of actions. To this end, we will design neurocognitive architectures that can segment perceived demonstration into meaningful units and relate them in a semantic way to objects on the table so that NICOL can perform the demonstrated actions from its perspective, even in a changed environment, e.g., with objects placed on the table differently. Both scenarios combine NICOL's motor abilities with social interaction capability. Furthermore, we will use NICOL as a platform for ongoing research for neural and neuro-hybrid IK and motion planning approaches for robots with a restricted, human-like range of motion.

## ACKNOWLEDGMENT

The authors would like to thank Phuong Nguyen for his contributions to the NICOL robot, Lennart Clasmeier for his work on the NICOL simulation, and Sergio Lanza for his feedback on the manuscript.

## REFERENCES

- [1] A. L. Thomaz, G. Hoffman, and C. Breazeal, "Reinforcement learning with human teachers: Understanding how people want to teach robots," in *Proc. 15th IEEE Int. Symp. Robot Hum. Interact. Commun. (ROMAN)*, Sep. 2006, pp. 352–357.
- [2] V. Srinivasan and L. Takayama, "Help me please: Robot politeness strategies for soliciting help from humans," in *Proc. CHI Conf. Hum. Factors Comput. Syst.*, May 2016, pp. 4945–4955.
- [3] M. Kerzel, T. Pekarek-Rosin, E. Strahl, S. Heinrich, and S. Wermter, "Teaching NICO how to grasp: An empirical study on crossmodal social interaction as a key factor for robots learning from humans," *Frontiers Neurobot.*, vol. 14, p. 28, Jun. 2020.
- [4] B. Mutlu, A. Terrell, and C.-M. Huang, "Coordination mechanisms in human-robot collaboration," in *Proc. Workshop Collaborative Manipulation, 8th ACM/IEEE Int. Conf. Hum.-Robot Interact.* Princeton, NJ, USA: Citeseer, Mar. 2013, pp. 1–6.
- [5] M. Kerzel, E. Strahl, S. Magg, N. Navarro-Guerrero, S. Heinrich, and S. Wermter, "NICO—Neuro-inspired companion: A developmental humanoid robot platform for multimodal interaction," in *Proc. 26th IEEE Int. Symp. Robot Hum. Interact. Commun. (RO-MAN)*, Aug. 2017, pp. 113–120.
- [6] J. Mišeikis, I. Brijacak, S. Yahyanejad, K. Glette, O. J. Elle, and J. Torrens, "Transfer learning for unseen robot detection and joint estimation on a multi-objective convolutional neural network," in *Proc. IEEE Int. Conf. Intell. Saf. for Robot. (ISR)*, Aug. 2018, pp. 337–342.
- [7] Universal Robots. *UR5 Technical Specifications*. Accessed: Dec. 27, 2020. [Online]. Available: <http://www.universal-robots.com>

- [8] R. Bischoff, J. Kurth, G. Schreiber, R. Koeppel, A. Albu-Schaeffer, A. Beyer, O. Eiberger, S. Haddadin, A. Stemmer, G. Grunwald, and G. Hirzinger, "The KUKA-DLR lightweight robot arm—A new reference platform for robotics research and manufacturing," in *Proc. 41st Int. Symp. Robot. (ISR), 6th German Conf. Robot. (ROBOTIK)*, Jun. 2010, pp. 1–8.
- [9] Franka Emika. (2022). *Franka Research 3*. Accessed: Dec. 27, 2022. [Online]. Available: <https://www.franka.de/research>
- [10] G. Metta, L. Natale, F. Nori, G. Sandini, D. Vernon, L. Fadiga, C. von Hofsten, K. Rosander, M. Lopes, J. Santos-Victor, A. Bernardino, and L. Montesano, "The iCub humanoid robot: An open-systems platform for research in cognitive development," *Neural Netw.*, vol. 23, nos. 8–9, pp. 1125–1134, Oct. 2010.
- [11] S. Al Moubayed, J. Beskow, G. Skantze, and B. Granström, "Furhat: A back-projected human-like robot head for multiparty human-machine interaction," in *Cognitive Behavioural Systems* (Lecture Notes in Computer Science), vol. 7403, A. Esposito, A. M. Esposito, A. Vinciarelli, R. Hoffmann, V. C. Müller, Eds. Berlin, Germany: Springer, 2012.
- [12] I. Ha, Y. Tamura, H. Asama, J. Han, and D. W. Hong, "Development of open humanoid platform DARwIn-OP," in *Proc. SICE Annu. Conf.*, Sep. 2011, pp. 2178–2181.
- [13] M. Schwarz et al., "Humanoid TeenSize open platform NimRo-OP," in *RoboCup 2013: Robot World Cup XVII* (Lecture Notes in Computer Science), vol. 8371, S. Behnke, M. Veloso, A. Visser, and R. Xiong, Eds. Berlin, Germany: Springer, 2014.
- [14] P. Allgeuer, H. Farazi, M. Schreiber, and S. Behnke, "Child-sized 3D printed igus humanoid open platform," in *Proc. IEEE-RAS 15th Int. Conf. Humanoid Robots (Humanoids)*, Seoul, South Korea, Nov. 2015, pp. 33–40.
- [15] G. Ficht, H. Farazi, A. Brandenburger, D. Rodriguez, D. Pavlichenko, P. Allgeuer, M. Hosseini, and S. Behnke, "NimRo-OP2X: Adult-sized open-source 3D printed humanoid robot," in *Proc. IEEE-RAS 18th Int. Conf. Humanoid Robots (Humanoids)*, Beijing, China, Nov. 2018, pp. 1–9.
- [16] A. K. Pandey and R. Gelin, "A mass-produced sociable humanoid robot: Pepper: The first machine of its kind," *IEEE Robot. Autom. Mag.*, vol. 25, no. 3, pp. 40–48, Sep. 2018.
- [17] M. Lapeyre, P. Rouanet, J. Grizou, S. N'Guyen, A. L. Falher, F. Depaetre, and P.-Y. Oudeyer, "Poppy: Open source 3D printed robot for experiments in developmental robotics," in *Proc. 4th Int. Conf. Develop. Learn. Epigenetic Robot.*, Oct. 2014, pp. 173–174.
- [18] O. Stasse, T. Flayols, R. Budhiraja, K. Giraud-Esclasse, J. Carpentier, J. Mirabel, A. D. Prete, P. Souères, N. Mansard, F. Lamiraux, J.-P. Laumond, L. Marchionni, H. Tome, and F. Ferro, "TALOS: A new humanoid research platform targeted for industrial applications," in *Proc. IEEE-RAS 17th Int. Conf. Humanoid Robot. (Humanoids)*, Nov. 2017, pp. 689–695.
- [19] T. Asfour, M. Waechter, L. Kaul, S. Rader, P. Weiner, S. Ottenhaus, R. Grimm, Y. Zhou, M. Grotz, and F. Paus, "ARMAR-6: A high-performance humanoid for human-robot collaboration in real-world scenarios," *IEEE Robot. Autom. Mag.*, vol. 26, no. 4, pp. 108–121, Dec. 2019.
- [20] E. Dean-Leon, B. Pierce, F. Bergner, P. Mittendorfer, K. Ramirez-Amaro, W. Burger, and G. Cheng, "TOMM: Tactile omnidirectional mobile manipulator," in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, May 2017, pp. 2441–2447.
- [21] *Robotics Automation for Warehousing, 3PLs, Distribution, Manufacturing, Fetch Robot.*, San Jose, CA, USA, 2019.
- [22] A. Parmiggiani, L. Fiorio, A. Scalzo, A. V. Sureshbabu, M. Randazzo, M. Maggiali, U. Pattacini, H. Lehmann, V. Tikhonoff, D. Domenichelli, A. Cardellino, P. Congiu, A. Pagnin, R. Cingolani, L. Natale, and G. Metta, "The design and validation of the R1 personal humanoid," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS)*, Sep. 2017, pp. 674–680.
- [23] N. Churamani, M. Kerzel, E. Strahl, P. Barros, and S. Wermter, "Teaching emotion expressions to a human companion robot using deep neural architectures," in *Proc. Int. Joint Conf. Neural Netw. (IJCNN)*, Anchorage, AK, USA, May 2017, pp. 627–634.
- [24] C. Bartneck, D. Kulić, E. Croft, and S. Zoghbi, "Measurement instruments for the anthropomorphism, animacy, likeability, perceived intelligence, and perceived safety of robots," *Int. J. Social Robot.*, vol. 1, no. 1, pp. 71–81, Jan. 2009.
- [25] N. Churamani, P. Barros, E. Strahl, and S. Wermter, "Learning empathy-driven emotion expressions using affective modulations," in *Proc. Int. Joint Conf. Neural Netw. (IJCNN)*, Jul. 2018, pp. 1400–1407.
- [26] H. G. Ng, P. Anton, M. Brügger, N. Churamani, E. Fließwasser, T. Hummel, J. Mayer, W. Mustafa, T. L. C. Nguyen, Q. Nguyen, M. Soll, S. Springenberg, S. Griffiths, S. Heinrich, N. Navarro-Guerrero, E. Strahl, J. Twiefel, C. Weber, and S. Wermter, "Hey robot, why don't you talk to me?" in *Proc. IEEE Int. Symp. Robot Hum. Interact. Commun. (RO-MAN)*, Aug. 2017, pp. 728–731.
- [27] H. B. Mohammadi, N. Xirakia, F. Abawi, I. Barykina, K. Chandran, G. Nair, C. Nguyen, D. Speck, T. Alpay, S. Griffiths, S. Heinrich, E. Strahl, C. Weber, and S. Wermter, "Designing a personality-driven robot for a human-robot interaction scenario," in *Proc. Int. Conf. Robot. Autom. (ICRA)*, Montreal, QC, Canada, May 2019, pp. 4317–4324.
- [28] M. Mori, K. F. MacDorman, and N. Kageki, "The uncanny valley [from the field]," *IEEE Robot. Autom. Mag.*, vol. 19, no. 2, pp. 98–100, Jun. 2012.
- [29] D. Coleman, I. Şucan, S. Chitta, and N. Correll, "Reducing the barrier to entry of complex robotic software: A MoveIt! Case study," *J. Softw. Eng. Robot.*, vol. 5, no. 1, pp. 3–16, May 2014.
- [30] R. Smits. *KDL: Kinematics and Dynamics Library*. Accessed: Oct. 6, 2023. [Online]. Available: <http://www.orocos.org/kdl>
- [31] P. Beeson and B. Ames, "TRAC-IK: An open-source library for improved solving of generic inverse kinematics," in *Proc. IEEE-RAS 15th Int. Conf. Humanoid Robots (Humanoids)*, Nov. 2015, pp. 928–935.
- [32] S. Starke, N. Hendrich, and J. Zhang, "A memetic evolutionary algorithm for real-time articulated kinematic motion," in *Proc. IEEE Congr. Evol. Comput. (CEC)*, Jun. 2017, pp. 2473–2479.
- [33] P. S. Ruppel, "Performance optimization and implementation of evolutionary inverse kinematics in ROS," M.S. thesis, Dept. Comput. Sci., Univ. Hamburg, Hamburg, Germany, 2017.
- [34] *BioIK Service: ROS Service for Calling BioIK From Python or Java*. Accessed: Oct. 6, 2023. [Online]. Available: [https://github.com/TAMS-Group/bio\\_ik\\_service](https://github.com/TAMS-Group/bio_ik_service)
- [35] M. Kerzel and S. Wermter, "Neural end-to-end self-learning of visuomotor skills by environment interaction," in *Artificial Neural Networks and Machine Learning—ICANN 2017* (Lecture Notes in Computer Science), vol. 10613, Cham, Switzerland: Springer, Sep. 2017, pp. 27–34.
- [36] S. Marsland, *Machine Learning: An Algorithmic Perspective*, 2nd ed. Boca Raton, FL, USA: CRC Press, 2014.
- [37] M. Kerzel, J. Spisak, E. Strahl, and S. Wermter, "Neuro-genetic visuomotor architecture for robotic grasping," in *Artificial Neural Networks and Machine Learning—ICANN 2020* (Lecture Notes in Computer Science), Bratislava, Slovakia: Springer, 2020, pp. 533–545.
- [38] J. Bergstra, D. Yamins, and D. Cox, "Making a science of model search: Hyperparameter optimization in hundreds of dimensions for vision architectures," in *Proc. 30th Int. Conf. Mach. Learn. (ICML)*, 2013, pp. 115–123.



**MATTHIAS KERZEL** received the M.Sc. and Ph.D. degrees in computer science from the University of Hamburg, Germany. He is currently a Postdoctoral Research and a Teaching Associate with the Knowledge Technology Group of Prof. Stefan Wermter, University of Hamburg. He has given lectures on knowledge processing in intelligent systems, neural networks, and bio-inspired artificial intelligence. He is involved in the international SFB/TRR-169 large-scale project on crossmodal learning. His research interests include developmental neurorobotics, hybrid neurosymbolic architectures, explainable AI, and human-robot interaction. He is a Secretary of the European Neural Network Society (ENNS) and involved in the organizing committee of the ICANN conferences.



**PHILIPP ALLGEUER** received the B.Eng./B.Sc. degree (Hons.) from The University of Adelaide, Australia, graduating with the overall top University Medal, and the Ph.D. degree in computer science from the University of Bonn, with a focus on humanoid robotics and bipedal locomotion. He is currently a Postdoctoral Research Associate with the Knowledge Technology Group, University of Hamburg. His research interests include making advances in deep learning methods and architectures, and uniting these with physical robot platforms to provide better human-robot interaction experiences and capabilities.



**ERIK STRAHL** received the degree in computer science and law from the Hamburg University of Applied Sciences and the University of Hamburg. He is currently a Technical Engineer with the Knowledge Technology Group, Department of Informatics, University of Hamburg. One focus of the work is the development and preparation of robots for the research of the Knowledge Technology Group on the software and hardware side. In this position, he participated in the development of the NICO and NICOL robots and their usage of them in many research projects.



**NICOLAS FRICK** received the B.Sc. degree in applied computer science from the University of Duisburg-Essen, with a focus on embedded systems. He is currently pursuing the M.Sc. degree in computer science with the University of Hamburg. He completed vocational training and was involved in the industrial sector as a Mechatronic Engineer before moving to academia. He is also a part of the NICOL Humanoid Robot Engineering Team. His research interest includes the applicability of deep learning methods to humanoid robotic grasping.



**JAN-GERRIT HABEKOST** received the B.Sc. degree in business informatics from the Carl von Ossietzky University of Oldenburg. He is currently pursuing the M.Sc. degree in computer science with the University of Hamburg. He is a Student Assistant with the Knowledge Technology Group, University of Hamburg. His research interests include humanoid robotics, deep learning, and neuro-robotic applications.



**MANFRED EPPE** is currently the Chief Engineer and the Head of the Robotics Laboratory, Institute for Data Science Foundations, Hamburg University of Technology. His current research is in cognitive robotics and machine learning. Previously, he was a Postdoctoral Research Associate with the Knowledge Technology Group, University of Hamburg. He has also held positions with the University of California at Berkeley; the Artificial Intelligence Research Institute, Barcelona; and the University of Bremen. He has published in numerous renowned venues, including Nature Machine Intelligence, Artificial Intelligence, and the International Conference on Intelligent Robots and Systems (IROS). He has coordinated several research projects funded by the German Research Foundation (DFG) and the Volkswagen Stiftung. His research interests include computational cognitive models of representation learning and reinforcement learning for simulated and physical agents and robots.



**STEFAN WERMTER** (Member, IEEE) is currently a Full Professor with the University of Hamburg, Germany, and the Director of the Knowledge Technology Group, Department of Informatics. He has previously held positions with the University of Dortmund, Dortmund, Germany; the University of Massachusetts, Amherst, MA, USA; the International Computer Science Institute, Berkeley, CA, USA; and the University of Sunderland, Sunderland, U.K. His main research interests include neural networks, hybrid knowledge technology, neuroscience-inspired computing, cognitive robotics, natural language processing, and human-robot interaction. He has been an Associate Editor of the IEEE TRANSACTIONS ON NEURAL NETWORKS AND LEARNING SYSTEMS. He is an Associate Editor of *Connection Science* and *International Journal for Hybrid Intelligent Systems*, and on the editorial board of the *Cognitive Computation* and *International Journal of Computational Intelligence Systems*. He is also a coordinator of the international doctoral training network TRAIL and co-coordinator of the International Collaborative Research Center on Crossmodal Learning (TRR-169). He is currently serving as the President of the European Neural Network Society.

...