IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received November 23, 2021, accepted December 23, 2021, date of publication January 5, 2022, date of current version January 12, 2022.

Digital Object Identifier 10.1109/ACCESS.2022.3140513

Deep Learning-Based Recognition and Analysis of Limb-Independent Dog Behavior for Ethorobotical Application

BALÁZS NAGY[1] AND PÉTER KORONDI[2], (Senior Member, IEEE)
[1]Department of Mechatronics, Optics and Mechanical Engineering Informatics, Budapest University of Technology and Economics, 1111 Budapest, Hungary
[2]Department of Mechatronics, Faculty of Engineering, Debrecen University, 4028 Debrecen, Hungary

Corresponding author: Balázs Nagy (nagybalazs@mogi.bme.hu)

This work was supported in part by the Budapest University of Technology and Economics (BME) through the National Research, Development and Innovation Office (NRDI) Fund based on the charter of bolster issued by the NRDI Office under the auspices of the Ministry for Innovation and Technology under Grant BME-NC (TKP2020 NC) and Grant BME-IE-MIFM (TKP2020 IES), and in part by the Hungarian Research Fund under Grant OTKA K12051.

This work involved human subjects or animals in its research. Approval of all ethical and experimental procedures and protocols was granted by the National Science and Research Ethics Committee (Hungary) (21/2015).

ABSTRACT This paper presents a Behavior Transfer System (BTS) to model the behavior patterns of dogs and make it possible to implement the behavior patterns on mobile robots. The system relies on an iSpace based measurement system and a deep learning prediction algorithm. With the help of the measurement system, ethological measurements can be automatized to eliminate human coding errors and make the data collection process more robust and consistent. The trained neural networks have a dual purpose. First, the neural networks can be utilized to analyze ethological measurements and predict different behavior patterns of the dog. Test results show that the implemented neural networks can effectively predict the attention of the dog with 88% accuracy, the tail waging with 82% accuracy, and the contact seeking behavior with 88% accuracy. Second, implementing the neural networks previously trained on dogs can serve as a robot operational behavioral model which mimics the behavior pattern of a dog after an adequate mathematical abstraction that maps the movements of the dog into a robot movement set. The presented method of this paper can be applied to automatize the behavior coding work of ethologists and the trained neural network can be used as an abstract robot behavior control module.

INDEX TERMS Ethology, ethorobotics, intelligent space, deep learning, behavior model, limb-independent behavior.

I. INTRODUCTION

In recent years automatization and robotics have been used in more and more areas besides the industrial sector. The rapid spreading of robots in daily living environments directs the attention to the field of human-robot interactions (HRI) [1]–[3], human-robot interaction models (MIHR) [4] and social behavior of robots [5]–[7]. The so-called service or social robots are becoming part of everyday life, and this coexistence between humans and robots raises some questions. How should a service robot look like? How should a social robot behave with humans? Will a robot pass Ainsworth's Strange Situation Test? [8] To develop an autonomous robot with proper appearance and movement set to be accepted by humans is a challenging task.

In the field of HRI, recent studies utilize digital twin [9] technology. A digital twin is a manifestation of a physical object in a virtual environment. A digital thread provides the connection between the physical object and twin, which means the "twin" is dependent on the digital thread to maintain accuracy. Digital twins are used for training and testing human-robot collaboration scenarios [10] without putting humans at risk and helpful tools to collect data and provide a better understanding of a specified task [11]. A digital twin-based simulation combined with a deep learning neural network results in an effective and accurate robot controller algorithm [12].

According to the uncanny valley theory of Masahiro [13] a response of a person to a humanlike robot would abruptly

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

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/

shift from empathy to revulsion as it approached, but failed to attain, a lifelike appearance. This descent into eeriness is known as the uncanny valley. Furthermore, the theory can be expanded into a more general form. Suppose a robot is shaped after a previously known living species, but the likeliness reaches a certain degree associated with the uncanny valley effect. In that case, it will fail to connect with humans. Thus a better approach is to treat robots as a unique and new species and develop their appearance and behavior according to it. However, the behavioral patterns can be modeled after well-known creatures like dogs. A wide group of social robots uses dogs as models [14]–[16] since dogs always have been reliable companions for humans. A robot design approach based on the behavior of dogs and other pet-like animals utilized in several applications [17], [18].

A. ETHOLOGICAL APPROACH

Ethorobotics [19] is a newly emerging interdisciplinary field of science that aims to combine robotics with ethologically inspired behavior models [20]. According to the methodology of ethorobotics instead of taking into account any concept of human or animal-like appearance the robot design should be based on the environment and the desired skill set of the robot [21]. Simple and optimal design with the capability of minimal social behavior. In the ethorobotic approach, embodiment and behavior have a strong functional relationship. In this way, the sub-assemblies of the robot serve a dual purpose. For example, a 2 DoF camera moving mechanism provide a wider observation view for the robot to examine the environment. Besides the desired functional ability, the exact mechanism can be used to express attention by focusing or following special items or persons near the robot. This multi-purpose attitude is essential from the view of ethorobotics.

Following the principles of ethorobotics numerous robot agents have been made. Vector [22] is one of the newest members of social robots. Designed by Anki Vector is an affordable robot companion and helper for people at home. The robot has its own personality which means it can perform various behavioral patterns such as greeting a familiar face or initiating playing sequences.

A more concrete approach of social robot design to use dogs as behavior model [14], [23]. A behavioral model of a pet dog is not so complex as a human psychological model, hence it is a better match for a smaller robot with limited computational capacity. The simpler cognitive model can be mapped to different robots which do not necessarily show similarities in appearance.

As a matter of functionalities guide dogs inspired some robots as well [24], [25]. Sunflower [15] is a dog-inspired robot that can express non-verbal communicative behaviors like attention-seeking, gaze alternation, and looking back. Special assistance dogs mean an excellent relief for disabled people. In the future, helper robots, like Sunflower, can provide a cheaper and reliable option instead of highly trained dogs. To achieve this goal, first we need to understand the behavior of dogs, and in the second phase, we need to implement the behavior patterns into robots successfully.

MogiRobi [18] is a partially dog alike robot which is implemented with an ecologically inspired fuzzy rule-based behavior model [26]. The model was tuned according to dog behavior experiments supervised by ethologists. A previous study [27] showed that dogs could be classified into seven different personality types. The implemented fuzzy model is capable of distinguishing these categories with the help of the tune-able parameters. In this case, the robot can be operated with different personalities. To achieve this level of complexity in the field of behavior models, numerous dogs were studied. The rules of the implemented fuzzy system were explicitly predefined according to the result of the studies.

B. DATA COLLECTION AND ANALYSIS

Tuning and evaluating the behavior models in the field of HRI emerged a need for a specific data set and a measurement system. To make the models more accurate, proper quality data is needed. In the field of ethology, using ethograms is a widespread and well-known method to code behavior data [28]–[30]. Ethograms contain the behavior patterns of a particular animal in a given time frame. The technique requires human scientists to decode sequences from a recorded videotape manually. Hence the method is prone to human errors and very time-consuming. Guiding principles are also made on adjusting a traditional ethogram to fit the need of autonomous robots [31] and how to determine observable input variables. The process is mainly made manually by a human scientist who analyses the video frame by frame. To make the data procession task faster and more reliable, deep learning-based solutions were implemented. DeepLabCut [32] is a widely-used open-source toolkit that can estimate the pose of different animals in an image. Labuguen *et al.* [33] used the DeepLabCut toolkit to estimate the pose of monkeys successfully. Fujimori *et al.* [34] implemented a cat behavior classification algorithm trained on images from a fixed camera. The Facial Action Coding System (FACS) is a system to taxonomize human facial movements by their appearance on the face [35]. There are multiple specialized FACS applications for different species [36]. The key point detection deep learning-based system also can be a solid base for problem-specific solutions. Andersen *et al.* [37] implemented a deep recurrent two-stream network for end-to-end detection of pain in the case of horses. The test result showed that the algorithm outperformed human scientists and achieved a 73.5% accuracy. However, the before-mentioned algorithms proved to be efficient. They mostly rely on image processing and focus on one particular species from the outside. Also, the before mentioned applications are limb-dependent and concentrate on a pose or posture of the examined species. A more abstract behavior model should consider the studied species in a broader environment instead of focusing on one individual. A possible solution is to take into account variables that describe the connection with other agents in the environment.

Vincze *et al.* [38] introduced a simulated environment that is specialized in ethologically inspired HRI observation and evaluation. The program is capable of exhibiting pre-programmed behaviors and automatically coding behavior elements. The system is prepared to accept raw environmental data. However, it was used only with simulated data. To extend the system to be used in a real environment a new and reliable sensory measurement system is needed. If a proper measurement system could provide a suitable raw data set the evaluation of real-life experiments becomes possible.

C. PROPOSED APPROACH

The new aspect in this paper is the limb-independent approach of a behavior pattern analysis of a natural agent, like a dog. Limbs are the key elements of animal motions in general to determine different behavior patterns, but limbs cannot be utilized effectively in the case of a wheeled robot. A more reliable way is to observe the spacing of the different agents in a multi-agent environment.

In our research, we would like to concentrate on such a complex measurement and analytical coding solution that can take into account interactions between different agents and describe basic behaviors—furthermore implementing a deep learning-based behavior coding algorithm. The input variables of the deep learning algorithm should be chosen to match the sensory equipment of a small autonomous robot. In this way, the algorithm can create a bridge between behavior analysis and real robot applications and can serve a dual purpose.

To minimize the possibility of human error and make the data collection process more efficient, we propose a new measurement system. The system is prepared to observe dogs and humans simultaneously in a specified room, focusing on individual behavior patterns and their interactions. Collect and store data, and based on the observations, the behavior analysis can be automatized. Using deep learning algorithms the system can learn how to predict a behavior which can be used to generate more precise ethograms to describe the behavior of the dog or the prediction capability can be used to implement a dog-inspired behavior model into an autonomous mobile robot in the future.

The paper is organized as follows. Section II summarize the motivation of the paper and highlights the main challenges. Section III describes the experimental setup used to examine dogs and human agents. Detailing the design of the behavior transfer system, the developed measurement system, methods, and the neural network-based prediction approach. Section IV presents the results of the prediction in the case of three different behavior patterns. Finally, Section V presents our conclusion and future plan.

II. PROBLEM STATEMENT

The harmonious cohabitation of humans and autonomous robots [39] is the ultimate goal of ethorobotics. To achieve the desired goal, social robots should be accepted by humans. This means that besides a functional skill set, robots have to maintain some behavior engine or model. Previous studies were made to discover the influencing variables for the acceptance of social robots [40] and make robots welcomed in everyday life [41]. Human behavior-based solutions often suffer from eeriness and indicate discomfort [42]. In this manner, social robots of the future should not copy human behavior more and more accurately and thus become a technical clone of a human, but appear as a new artificial species in the human environment. However, behavior elements of a robot social skillset can mimic simple animal behaviors. Modeling animal-like behavior patterns into a mobile robot has some practical advantages. First of all, the complexity of a pet creature is more suitable for a small-scale autonomous robot and does not generate irrational expectations while interacting with humans. On the other hand, pets, especially dogs, have already proved to be good companions and well integrated into the human environment. Many people believe that the dog is the most loyal companion of a man, so dogs can serve as an excellent model for the new artificial species that serve humans. Modeling robot behavior after dogs also implies that the interaction between a robot and a human also should represent an inter-species interaction, like for example a dog and owner relationship rather than a human-human type interaction. To develop and implement a dog behavior-based behavior model on a mobile robot emerges a need for a behavior transfer system (BTS). A system that is capable of observing a dog skill set and transferring it into a robot on an abstract level. The goal is to use dogs as base inspiration for behavior patterns but not as an exact copy. While the digital twin concept depends on the high-level similarity between the physical agent and the model, the purpose of the BTS is to remap similar patterns between two different agents. Dogs use legs for moving, robots move on wheels but the positioning strategies of both agents can be matched.

Theoretically, the task is to observe a behavior of a dog in different situations and the interactions of the dog with humans. With enough data, the desired behavior patterns could be learned by a machine learning algorithm. However, in practice, a task is way more complex. A fundamental difference presumably persists for a long time between pet animals and commercial serving robots. Walking robots are much more expensive than wheeled robots, so commercial robots that can be sold in large numbers are expected to be wheeled for a long time. Bearing in mind this consideration a new level of abstraction should be incorporated into standard ethological observation, where instead of limb movements, other elements of behavior should be observed. This justifies the development of a behavior transfer system for recognizing particular limb-independent behavior patterns based on the intelligent space concept described in the article.

The methodology of the behavior transfer system requires three key elements: an ethologically approved test, a proper measurement system, and a learning algorithm. Interactions between species have been studied by ethnologists and a number of methods like cluster analysis [43] or multi-layer network analysis [44] have been developed for this purpose. To keep the process robust and repeatable, a well-described and controlled experiment scenario is used, called the Strange Situation Test by Ainsworth. The original test [45] was designed to examine the bond between a baby and a caregiver. The test later was used to examine the interactions between a dog and the owner of the dog [27]. In our case, the future goal is to take one more step and implement the test to examine a human-robot interaction in case of a successful behavior transfer implementation. The test consists of seven different blocks. Each block is two minutes long and mainly starts with a passive phase and ends with a more active phase. The test provides versatile situations in a controlled sequence and environment thus, it is suitable to observe a relative complex behavior in the case of a dog-human interaction.

To collect proper quality and quantity data for a deep leaning based behavior analyzer is mandatory. There is no such available database in the literature containing numerical measurements about dog behaviors focusing on human interactions. To fulfill this need a novel measurement system and setup is needed. So far, the experiments were recorded on video, and the labeling of the data was made manually. The labeling of the data only includes predefined behavior patterns without the sensory measurement values. Processing data is cumbersome and laborious. To speed up the process and eliminate human error, an intelligent space-based measurement system is proposed. The system makes it possible to collect position, orientation, and other data automatically and controls the episodes without any external human help.

The final component of the behavior transfer system is the deep learning algorithm. During the learning phase, the algorithm is trained with a supervised learning method. In this case, the algorithm can learn how to decode automatically behavior patterns on the previously logged data by the measurement system. In this form in later, the algorithm can speed up the labeling and replace humans in the process. On the other hand, the trained algorithm can be used in a mobile robot as a behavior-based decision-making algorithm. Previously mentioned applications [33]–[37] utilized deep learning-based solutions on video stream data to observe and analyze animal behaviors, but these applications rely on image data only. Image-based labeling of the behavior patterns can be applied to analyze individual animals, but the decoded database does not contain numerical information about the cause of the observed behavior.

The main challenges that the paper addresses are the following. The first is to propose a dedicated measurement system to observe and analyze dog behaviors without human inference. The only requirement of the measurement system is the attached marker set to the observed agent. The marker sets make it possible to track variables that could play a significant role in triggering different behavior patterns in the case of dogs interacting with multiple humans. Since there is no such database available in the literature that could provide a solid base for deep learning solutions, the measurement system plays a key role. The observed variables are chosen

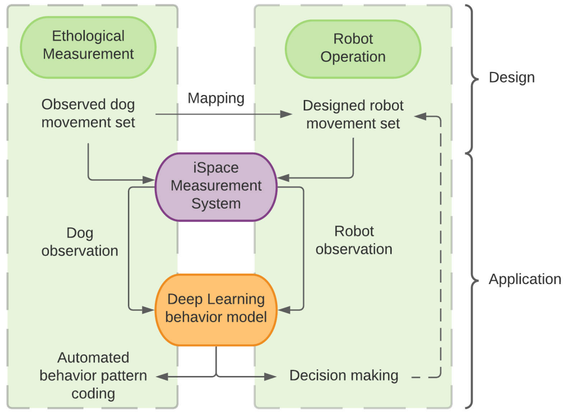


FIGURE 1. Behavior transfer flow chart.

so that an autonomous mobile robot onboard sensor system could reproduce similar signals in the future. For example, a LIDAR sensor provides distance measurement to calculate the relative distances between the agents. The second challenge is to implement a deep learning architecture that is capable of predicting behavior of a dog. In the short run, neural networks can help the work of human ethologists by automatizing the measurement decoding step and reducing human errors. In the long run, if the predictions are precise and the trained model is complex enough and similar to real dog behavior, the model can control an autonomous robot after an appropriate mapping. In this way, a dog mimicking behavior model implemented on the robot could help the social acceptance of the robots in everyday life.

III. BEHAVIOR TRANSFER SYSTEM (BTS)

This section describes the novel behavior Transfer System (BTS) for dog behavior coding and robot control. The schematic of the system is shown in Fig. 1.

A. DESIGN

The dual purpose of the method is to fulfill the requirements of the ethological measurements and a robot operating system in the same architecture. The design steps and application steps can be applied parallel. However, there are some key points in which this method differs from the traditional ethological approach. Since the main goal is to learn a behavior pattern of a dog and transfer them into an autonomous robot, an abstract mathematical mapping step is required. The simultaneous robot design makes it possible to mount the robot with adequate equipment to meet certain expectations. For example, feedback LED blinking frequency could mimic the tail wagging of a dog. Other key features are the observed movement sets. Traditional ethological pattern coding methods observe the posture of the dog or the limbs of the animal. Using mostly wheeled robots in the field of social robots, a sophisticated engineering point of view emerged. In robot control, the trajectory of the movement is more important since most of the robots do not have limbs. A limb independent behavior model contains mapping rules

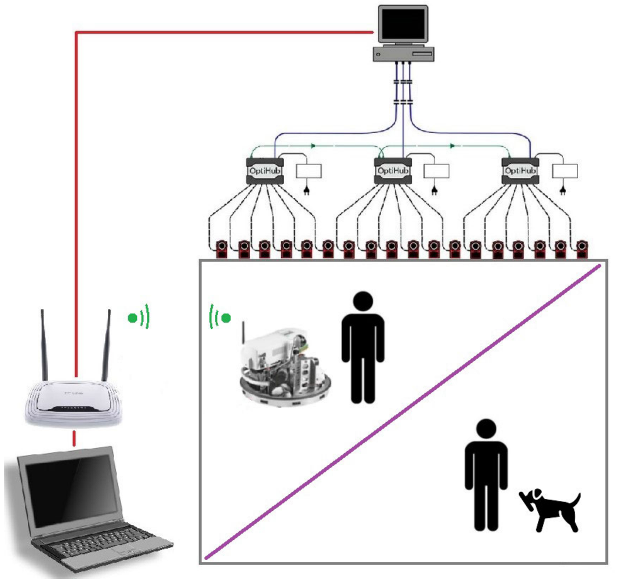


FIGURE 2. Intelligent space setup to track specified marker sets and log data position and orientation data or to provide sufficient data stream to control a robot.

to convert observable dog movements into possible robot feedback patterns. In this way, even a four-legged animal movement set could be transferred into a wheeled robot.

B. MEASUREMENT SYSTEM

An intelligent space (iSpace) concept [46] was used as a measurement system. The setup contains an OptiTrack system with 18 infrared cameras mounted in a room and using the Motive 3.10 software [47]. The observation space is roughly 5 m by 2.5 m. The schematic of the iSpace is shown in Fig 2.

The mounted cameras as a motion capture system can track infra reflective markers and marker sets. If at least three cameras see a marker at the same time the position of the spherical marker can be calculated. Three or more markers can be defined as a rigid body. As long as all of the rigid body markers are tracked, the 3D position and orientation of the rigid body can be calculated and obtained from the OptiTrack system. The position tracking error of the system is approximately 0.2 mm. Each different moving agents in the iSpace have a unique marker set, and the system tracks the position and orientation of each agent. Fig. 3 shows two marker sets in the case of a human holding a toy. The offset pivot point of the toy marker set matches the center of mass of the ball. Using this design, the system can track the ball even when it is grabbed. Tracked data can be logged or streamed via WiFi connection if an autonomous robot agent is inside the observed space.

In the first scenario, the iSpace setup was used to observe the behavior of a dog. The ethological measurement design focuses on the movements of a dog while interacting with humans. For the experiment, the strange situation test was used. The original test was introduced by Ainsworth to examine the attachment between a baby and a caregiver. Later the modified version of the test was used by ethologists to

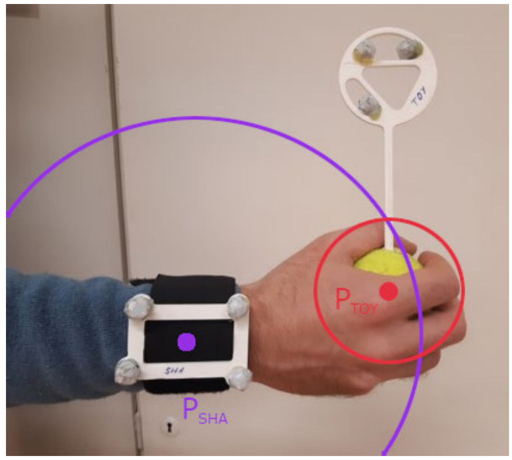


FIGURE 3. SHA and TOY marker sets and the calculated pivot points respectively P_{SHA} and P_{TOY}.

TABLE 1. Marker set abbreviations and references.

Name	Marker set reference
DOG	dog
OWN	owner of the dog
OHA	owner hand
STR	stranger to the dog
SHA	stranger hand
TOY	toy (tennis ball)
DOOR	door to the room

examine the attachment behavior of dogs. The measurement scenario contains seven episodes, each episode is two minutes long. In this case, the measurement is well defined and repeatable which makes it possible to examine and identify variables that can cause the different dog behaviors. The robot operation side needs to rely on variables that can be identified and measured in the case of a robot. The human distance from a robot or dog or the relative orientation of the agents is a suitable choice. Table 1. shows the marker sets used in the measurement.

The main goal was to collect enough quantitative data to teach a neural network. The position and orientation of the DOG, STR, and OWN marker set were recorded. Only the position data of the TOY, SHA, and OHA marker sets were tracked since the orientation ofthese marker sets are negligible. The orientation of the DOOR marker set was tracked. The OptiTrack system uses an absolute coordinate system. To make the dating uniform, a data preprocessing step was made to calculate the relative distances and angles of the observed marker sets compared to the dog. The representation of the calculated variables can be seen in Fig. 4. The relative distance and angular values respectively to the dog are calculated according to Eq. (1-8).

$$d_{OWN} = |\vec{P}_{OWN} - \vec{P}_{DOG}| \tag{1}$$

$$d_{STR} = |\vec{P}_{STR} - \vec{P}_{DOG}| \tag{2}$$

$$d_{TOY} = |\vec{P}_{TOY} - \vec{P}_{DOG}| \tag{3}$$

$$d_{DOOR} = |\vec{P}_{DOOR} - \vec{P}_{DOG}| \tag{4}$$

$$\theta_{TOY} = \theta_{DOG} - \arctan\left(\frac{(\vec{P}_{TOY} - \vec{P}_{DOG})_y}{(\vec{P}_{TOY} - \vec{P}_{DOG})_x}\right) \tag{5}$$

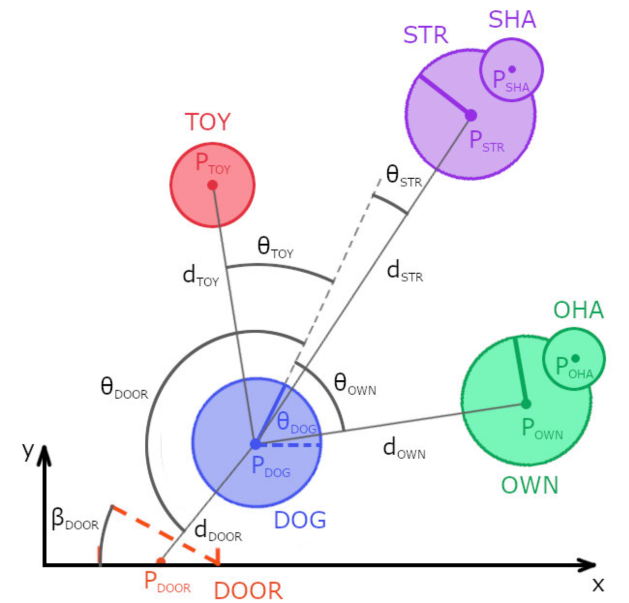


FIGURE 4. Schematic 2D representation of the specified marker sets in the absolute coordinate system of the OptiTrack observation space.

$$\theta_{STR} = \theta_{DOG} - \arctan\left(\frac{(\vec{P}_{STR} - \vec{P}_{DOG})_y}{(\vec{P}_{STR} - \vec{P}_{DOG})_x}\right) \quad (6)$$

$$\theta_{OWN} = \theta_{DOG} - \arctan\left(\frac{(\vec{P}_{OWN} - \vec{P}_{DOG})_y}{(\vec{P}_{OWN} - \vec{P}_{DOG})_x}\right) \quad (7)$$

$$\theta_{DOOR} = \theta_{DOG} - \arctan\left(\frac{(\vec{P}_{DOOR} - \vec{P}_{DOG})_y}{(\vec{P}_{DOOR} - \vec{P}_{DOG})_x}\right) \quad (8)$$

The TOY marker set was designed to make it possible to track the TOY even if it is partially hidden or in the mouth of a dog. The OHA, SHA, and DOOR marker sets have a supporting role. Hands markers were used to identify the ownership of the TOY. OHA and SHA were attached to the dominant hand of the owner and stranger, respectively. Before the measurement, the human participant was asked to use only their dominant hand when initiating playing with the dog. Based on threshold values and synchronous movements, the toy has four states: carried by the owner, carried by the dog, carried by the stranger, or not carried. The variables describing the ownership of the TOY calculated with Eq. (9-11).

$$STR\_has\_TOY = \begin{cases} 1, & \text{if } |\vec{P}_{TOY} - \vec{P}_{SHA}| \leq T_{STR} \\ 0, & \text{if } |\vec{P}_{TOY} - \vec{P}_{SHA}| > T_{STR} \end{cases} \quad (9)$$

$$OWN\_has\_TOY = \begin{cases} 1, & \text{if } |\vec{P}_{TOY} - \vec{P}_{OHA}| \leq T_{OWN} \\ 0, & \text{if } |\vec{P}_{TOY} - \vec{P}_{OHA}| > T_{OWN} \end{cases} \quad (10)$$

$$DOG\_has\_TOY = \begin{cases} 1, & \text{if } |\vec{P}_{TOY} - \vec{P}_{DOG}| \leq T_{DOG} \\ 0, & \text{if } |\vec{P}_{TOY} - \vec{P}_{DOG}| > T_{DOG} \end{cases} \quad (11)$$

$T_{STR} = 100$ mm, $T_{OWN} = 100$ mm, and $T_{DOG} = 150$ mm are predefined threshold values.

The DOOR also has a special role. The basic behavior pattern of the dogs shows that in the absence of the owner, most dogs sit by the door or the opening of the door attracts the attention of the dog. The opening of the door was calculated from the angle of the DOOR marker set according to Eq. (12).

TABLE 2. Episode details.

No.	Duration	Participants
1	2 min	DOG, OWN
2	2 min	DOG, OWN, STR
3	2 min	DOG, STR
4	2 min	DOG, OWN
5	2 min	DOG
6	2 min	DOG, STR
7	2 min	DOG, OWN

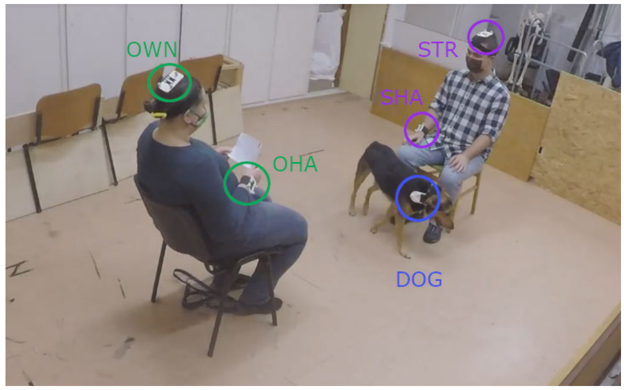


FIGURE 5. Ainsworth's Strange situation test episode 2. The owner on the left side is marked by the OWN and OHA marker sets. The dog is marked by the DOG marker set. The stranger on the right side is marked by the SHA and STR marker sets.

$$DOOR\_is\_OPEN = \begin{cases} 1, & \text{if } \beta_{DOOR} \geq T_{DOOR} \\ 0, & \text{if } \beta_{DOOR} < T_{DOOR} \end{cases} \quad (12)$$

$T_{DOOR} = 5°$ is a predefined threshold value.

The environment setup could vary depending on the room layout, and setting up a marker set on the door makes it easier to change the room setup. Furthermore, the door opening can help synchronize the episodes. The iSpace was programmed to execute the full measurement scenario without any external human interaction. The iSpace was mounted with a soundbar and a monitor. With the help of these outputs, it can instruct the human participants in the measurement to execute different tasks or episodes. Since the episode changes took some time, the transfer times were taken into account, and the DOOR marker set was used to identify the new episode. Every episode starts when the desired person leaves the room or arrives in the room and shuts the door. Table 2. shows the duration of the episodes and the observed participants.

Each episode started with a one minute-long passive phase when the owner and/or the stranger were instructed to sit still on the previously designated chair. In case the dog approached the owner or the stranger, they were allowed to pet the dog. Fig. 5 shows a passive phase state. After the passive phase in the second phase, the humans were instructed to seek contact with the dog actively and initiate playing.

C. DEEP LEARNING BEHAVIOR MODEL

Deep learning algorithms are used in a wide range of applications from controlling tasks [48], [49] to measurement system

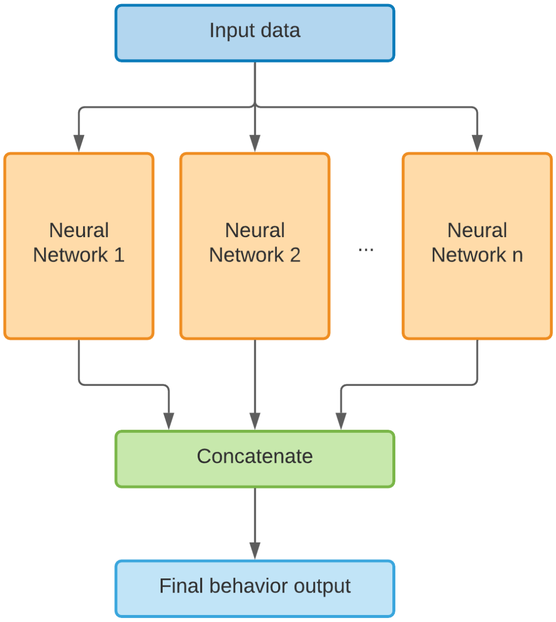


FIGURE 6. Behavior learning architecture.

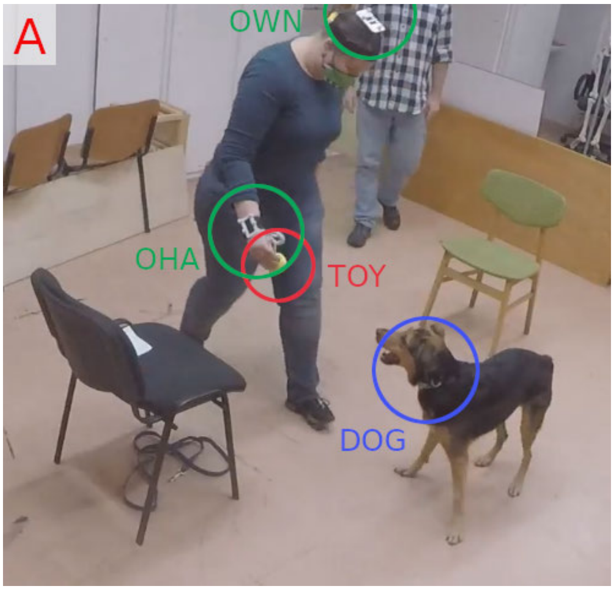


FIGURE 7. Mispredicted test case. The attention of the dog is on the toy, according to the human ethologist labeling the video. The deep learning algorithm predicted the owner as the focus of the attention.

TABLE 3. Accuracy comparison on different data sets.

Pattern	Training	Validation	Test
Contact	99%	92%	88%
Tail wag	94%	88%	82%
Attention	96%	74%	88%

optimization [50]. Neural networks can provide a flexible architecture with the capability to learn from structured data.

Ensemble learning [51] principles were used to create the architecture of the deep learning behavior model shown in Fig. 6. The input data were fed into separate neural network blocks. Each neural network is responsible for a specific behavioral pattern. In this setup, the model can handle multiple behavior patterns simultaneously, and the architecture is easily expandable in case of an integration of a new pattern. Multi-layer, fully connected feed-forward neural networks were used to prevent overfitting and preserve the computational cost efficiency of the algorithm. The neural networks are implemented in Python using Tensorflow Keras framework. During the training phase, an NVIDIA GeForce RTX 2080Ti graphic card was used. However, the neural network architecture was developed in a way that a Raspberry Pi 3 B+ could run the prediction step in the future.

The input parameters of the neural network were the distance of the agents from the dog and their relative orientation measured from the direction of the dog. Additional preprocessed input variables were calculated to determine if a marker set is tracked or not, the door is open or closed, and the ownership of the toy. In general, the input feature vector contains 19 variables. Tagged data records are treated as different instances and not as complex time series. To detect and avoid overfitting, the collected data is separated into training, validation, and test data sets. This is done by breaking the chronologically ordered data into smaller segments. The first 60% of a segment will be training data, and the other 20%-20% will be validation and test data, respectively. To ensure that each group gets data points from every episode. The number of segments should be divisible by the number of episodes.

In this paper three different output behavior variables were implemented and tested. The three examined behavior patterns of a dog are the tail wagging, the focus of the attention of the dog, and the physical contact between the dog and humans. The hyperparameters of the neural networks were set based on preliminary tests. A different neural network was assigned to each behavior pattern according to previous test results. The detailed neural networks used in the ensemble learning architecture shown in Fig. 6. are as follows. The first neural network is the tail wag predictor neural network which contains ten hidden layers with respectively 50-50-50-50-50-25-25-25-10-10 neurons in each layer. The second neural network is the contact predictor neural network which contains eight hidden layers with respectively 100-100-50-50-25-25-10-10 neurons in each layer. The third neural network is the attention predictor neural network which contains nine hidden layers with respectively 50-50-25-25-25-10-10-10-10 neurons in each layer. The implemented neural networks were trained with the Adam optimization algorithm. The combined predicted output of the neural networks presents the final behavior output. Depending on the application case, the behavior output could be the labeling of an ethological measurement video recording or the desired behavior of a mobile robot.

The neural networks were trained separately in the case of the different behavioral patterns. The training took 35 epochs, and the training metrics can be seen in Table. 3. During the training part, the early stopping method [52] was used to prevent overfitting and stop the training method at the best configuration.

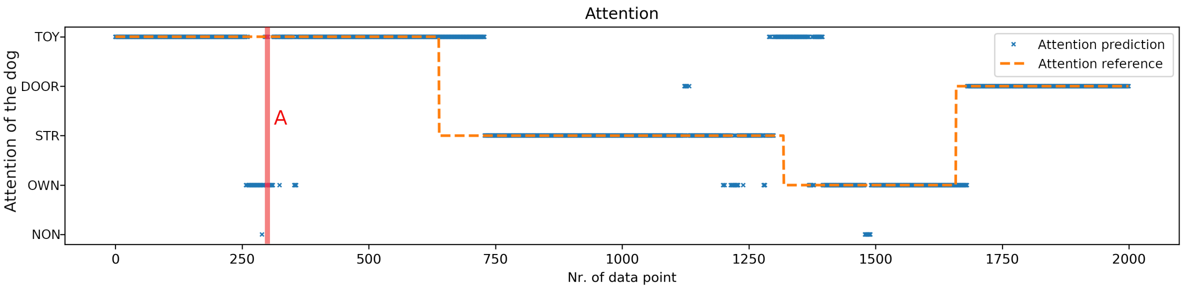


FIGURE 8. Result of attention prediction. (Dog looking at NON: non-specified location, OWN: owner, STR: stranger, DOOR: door, TOY: toy) The orange dash line represents the reference values decoded by a human ethologist. The blue dots represent the prediction of the proposed deep learning algorithm.

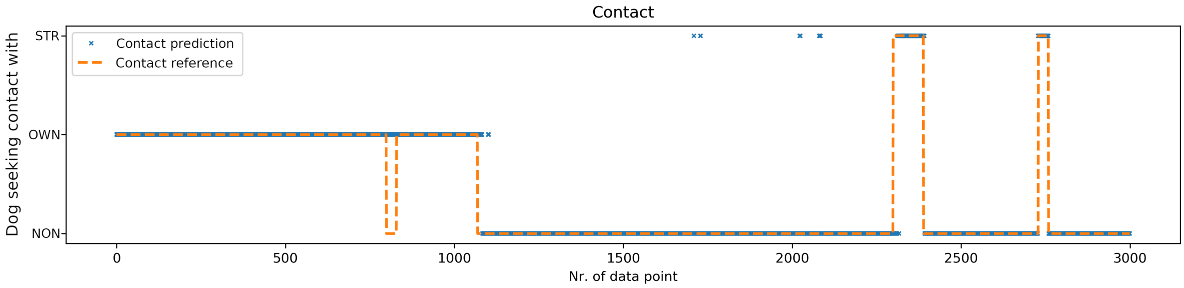


FIGURE 9. Result of contact prediction. (NON: No contact, OWN: Contact with the owner, STR: Contact with a stranger) The orange dash line represents the reference values decoded by a human ethologist. The blue dots represent the prediction of the proposed deep learning algorithm.

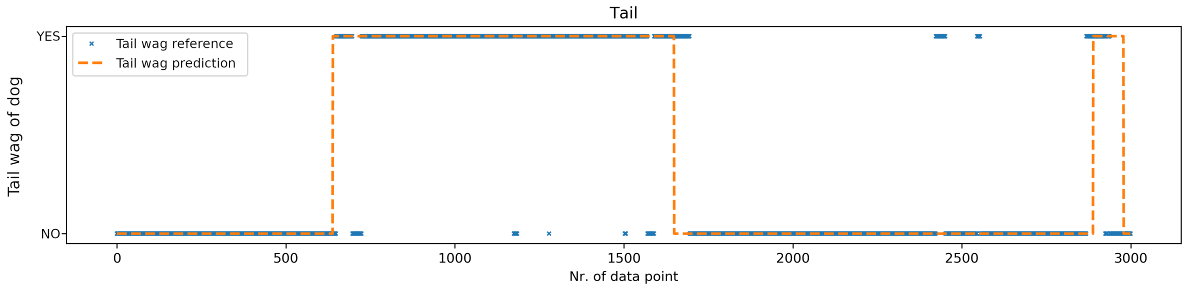


FIGURE 10. Result of tail wag prediction. (NO: No tail wag, YES: Tail wag) The orange dash line represents the reference values decoded by a human ethologist. The blue dots represent the prediction of the proposed deep learning algorithm.

IV. RESULTS

The National Science and Research Ethics Committee (Hungary) (21/2015) approved the experiment in which we examined the behavior of companion dogs with their owners. The participation was voluntary and the owners were informed about the main aim of the study and that they are allowed to interrupt the test at any time if needed. During the experiment three dog were examined and six person took part in the experiment. Each dog was measured multiple times. In this case, the dogs were tested in different setups, and separated data sets were logged for training and validating the neural network. The results stated in this paragraph refer to one dog. In this case an approximately 15 minutes long test was carried out including all seven episodes from the Ainsworth Strange Situation Test. The sampling time of the measurement system was set to 0.02 s, and 47590 data points were recorded in total. The entire database was separated into three different groups, so the training set contains 28518 data points, the validation set contains 9492 data points, and the test set contains 9580 data points. Each separate data set contains segments from all seven episodes. On this data set the neural network reached 88% accuracy in case of contact seeking prediction, 82% accuracy in case of tail wag prediction, and 88% accuracy in case of attention-seeking prediction. The compared accuracy of the three behavioral pattern training can be seen in Table 3.

The test set was a separate measurement not used in the training or validation set at all. The result of the attention prediction can be seen in Fig. 8. The result of the contact-seeking prediction can be seen in Fig. 9. The result of the tail wag prediction can be seen in Fig. 10. The orange line represents the reference signals decoded from a video by human scientists with Solomon coder. The blue dots represent the predicted signals from the neural network.

The results show that the errors made by the algorithm generate outliers. These prediction errors can be dealt with by implying a low pass filter after the prediction layer of the neural network. In this way, a more smooth signal could be achieved. Another type of error originates from human supervision. The reference signal was made with Solomon

Coder [53] by the human scientist. There are some cases in which it is very challenging to predict the correct answer. For example, when the owner is holding the toy, the attention of the dog is hard to decide. The situation can be seen in Fig. 7 and the exact location is marked by "A" on Fig. 8 between the 250th and 300th data point.

V. CONCLUSION

In this research, a novel measurement setup was proposed to examine the behavior of dogs and their interactions with humans. The motion capture camera system-based intelligent space was developed to collect quantitative data from the agents in the observed space tagged with infra reflective marker sets. The system can navigate through a predefined measurement scenario signaling the human participants the next step. During the measurement the system logs the orientation and position of all participants, and some calculated values to produce the necessary database needed by the proposed process. The collected numeric data and asynchronous video provide a good base for a supervised learning task. As a reference, human ethology scientists decode the behavior elements of a dog according to the video, which can be used as supervised labels during the training of the neural networks. The implemented neural network architecture can be used to predict different dog behaviors. This serves a dual purpose. On the one hand, the system can label the recorded video automatically. In this case, makes the work of ethology scientists faster and more reliable. Trained neural networks make the labeling process consistent meanwhile the human error can be minimized or neglected due to the system. Manual coding carried out by humans is always subjective while a trained system is consistently objective. On the other hand, special combined neural networks can be used to mimic dog behaviors more realistically. The learned dog behavioral patterns can be implemented in a real-life autonomous robot. The movement set of the robot can be modeled after a real dog. The method helps to develop socially more acceptable robot behaviors.

However, the system proved to be useful more future work will be required. More measurements with more dogs are needed to avoid overfitting and to build a database on which more generalized behaviors could be trained. The trained neural networks reliability could be higher with more data. The diversity in the examined dog could mean that different personality types of dogs can be targeted and implemented in a real-life application. Since the method in this article used a special data frame to work with recorded in a special test scenario more validation tests are needed. There is no dog behavior model available in the literature to implement or use it on a mobile robot to compare our achievements. Future work targets a test where a fully autonomous robot can interact with humans with the trained neural network-based behavior control system. Measuring the reactions of the human agents in the presence of a robot and comparing them to the measurements made with a dog can validate the success of the behavior element transfer.

ACKNOWLEDGMENT

The authors would like to thank Dr. János Botzheim for his insightful suggestions and proofreading of the manuscript, Soma Deme for the technical support during the measurement, and András Zöllner for helping with the measurement data preprocessing.

REFERENCES

[1] D. Herrera, F. Roberti, M. Toibero, and R. Carelli, "Human-robot interaction: Legible behavior rules in passing and crossing events," *IEEE Latin Amer. Trans.*, vol. 14, no. 6, pp. 2644–2650, Jun. 2016.

[2] H. Modares, I. Ranatunga, F. L. Lewis, and D. O. Popa, "Optimized assistive human–robot interaction using reinforcement learning," *IEEE Trans. Cybern.*, vol. 46, no. 3, pp. 655–667, Mar. 2016.

[3] D. Tang, B. Yusuf, J. Botzheim, N. Kubota, and C. S. Chan, "A novel multimodal communication framework using robot partner for aging population," *Expert Syst. Appl.*, vol. 42, no. 9, pp. 4540–4555, Jun. 2015.

[4] J. Perez, J. Aguilar, and E. Dapena, "MIHR: A human-robot interaction model," *IEEE Latin Amer. Trans.*, vol. 18, no. 9, pp. 1521–1529, Sep. 2020.

[5] P. Liu, D. F. Glas, T. Kanda, and H. Ishiguro, "Data-driven HRI: Learning social behaviors by example from human–human interaction," *IEEE Trans. Robot.*, vol. 32, no. 4, pp. 988–1008, Aug. 2016.

[6] M. Doering, D. F. Glas, and H. Ishiguro, "Modeling interaction structure for robot imitation learning of human social behavior," *IEEE Trans. Human-Mach. Syst.*, vol. 49, no. 3, pp. 219–231, Jun. 2019.

[7] J. Woo, J. Botzheim, and N. Kubota, "A socially interactive robot partner using content-based conversation system for information support," *J. Adv. Comput. Intell. Intell. Informat.*, vol. 22, no. 6, pp. 989–997, Oct. 2018.

[8] F. Kaplan, "Artificial attachment: Will a robot ever pass Ainsworth's strange situation test?" in *Proc. Humanoids IEEE-RAS Int. Conf. Humanoid Robots*, Jan. 2001, pp. 106–109.

[9] D. Gelernter, *Mirror Worlds or the Day Software Puts the Universe in a Shoebox: How Will it Happen and What it Will Mean*. London, U.K.: Oxford Univ. Press, 1991.

[10] E. Pairet, P. Ardon, X. Liu, J. Lopes, H. Hastie, and K. S. Lohan, "A digital twin for human-robot interaction," in *Proc. 14th ACM/IEEE Int. Conf. Hum.-Robot Interact. (HRI)*, Mar. 2019, p. 372.

[11] Q. Wang, W. Jiao, P. Wang, and Y. Zhang, "Digital twin for human-robot interactive welding and welder behavior analysis," *IEEE/CAA J. Automatica Sinica*, vol. 8, no. 2, pp. 334–343, Feb. 2021.

[12] K.-B. Park, S. H. Choi, J. Y. Lee, Y. Ghasemi, M. Mohammed, and H. Jeong, "Hands-free human–robot interaction using multimodal gestures and deep learning in wearable mixed reality," *IEEE Access*, vol. 9, pp. 55448–55464, 2021.

[13] M. Masahiro, "Bukimi no tani the uncanny valley," *Energy*, vol. 7, no. 4, pp. 33–35, 1970.

[14] G. Lakatos, "Dogs as behavior models for companion robots: How can human–dog interactions assist social robotics?" *IEEE Trans. Cognit. Develop. Syst.*, vol. 9, no. 3, pp. 234–240, Sep. 2017.

[15] K. L. Koay, G. Lakatos, D. S. Syrdal, M. Gacsi, B. Bereczky, K. Dautenhahn, A. Miklosi, and M. L. Walters, "Hey! There is someone at your door. A hearing robot using visual communication signals of hearing dogs to communicate intent," in *Proc. IEEE Symp. Artif. Life (ALife)*, Apr. 2013, pp. 90–97.

[16] M. Fujita, "On activating human communications with pet-type robot AIBO," *Proc. IEEE*, vol. 92, no. 11, pp. 1804–1813, Nov. 2004.

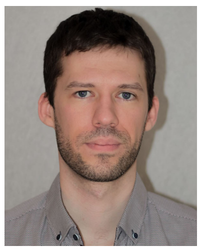
[17] T. Shibata, "Therapeutic seal robot as biofeedback medical device: Qualitative and quantitative evaluations of robot therapy in dementia care," *Proc. IEEE*, vol. 100, no. 8, pp. 2527–2538, Aug. 2012.

[18] G. Lakatos, M. Gácsi, V. Konok, I. Brúder, B. Bereczky, P. Korondi, and A. Miklósi, "Emotion attribution to a non-humanoid robot in different social situations," *PLoS ONE*, vol. 9, no. 12, pp. 1–32, 2014.

[19] Á. Miklósi, P. Korondi, V. Matellán, and M. Gácsi, "Ethorobotics: A new approach to human-robot relationship," *Frontiers Psychol.*, vol. 8, pp. 1–8, Jun. 2017.

[20] R. Numakunai, T. Ichikawa, M. Gacsi, P. Korondi, H. Hashimoto, and M. Niitsuma, "Exploratory behavior in ethologically inspired robot behavioral model," in *Proc. IEEE RO-MAN, 21st IEEE Int. Symp. Robot Hum. Interact. Commun.*, Sep. 2012, pp. 577–582.

[21] P. Korondi, B. Korcsok, S. Kovács, and M. Niitsuma, "Etho-robotics: What kind of behaviour can we learn from the animals?" *IFAC-PapersOnLine*, vol. 48, no. 19, pp. 244–255, 2015.
[22] *Vector Homepage*. Accessed: Jul. 20, 2021. [Online]. Available: http://www.digitaldreamlabs.com
[23] K. Ehsani, H. Bagherinezhad, J. Redmon, R. Mottaghi, and A. Farhadi, "Who let the dogs out? Modeling dog behavior from visual data," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 4051–4060.
[24] K. A. Hamed, V. R. Kamidi, W.-L. Ma, A. Leonessa, and A. D. Ames, "Hierarchical and safe motion control for cooperative locomotion of robotic guide dogs and humans: A hybrid systems approach," *IEEE Robot. Autom. Lett.*, vol. 5, no. 1, pp. 56–63, Jan. 2020.
[25] S. Tachi, K. Tanie, K. Komoriya, and M. Abe, "Electrocutaneous communication in a guide dog robot (MELDOG)," *IEEE Trans. Biomed. Eng.*, vol. BME-32, no. 7, pp. 461–469, Jul. 1985.
[26] S. Kovacs, D. Vincze, M. Gacsi, A. Miklosi, and P. Korondi, "Ethologically inspired robot behavior implementation," in *Proc. 4th Int. Conf. Human Syst. Interact. (HSI)*, May 2011, pp. 64–69.
[27] J. Topál, Á. Miklósi, V. Csányi, and A. Dóka, "Attachment behavior in dogs *(Canis familiaris)*: A new application of Ainsworth's (1969) strange situation test," *J. Comparative Psychol.*, vol. 112, no. 3, pp. 219–229, 1998.
[28] D. R. MacNulty, L. D. Mech, and D. W. Smith, "A proposed ethogram of large-carnivore predatory behavior, exemplified by the wolf," *J. Mammal.*, vol. 88, no. 3, pp. 595–605, Jun. 2007.
[29] L. A. Stanton, M. S. Sullivan, and J. M. Fazio, "A standardized ethogram for the felidae: A tool for behavioral researchers," *Appl. Animal Behav. Sci.*, vol. 173, pp. 3–16, Dec. 2015.
[30] C. Hall and C. Heleski, "The role of the ethogram in equitation science," *Appl. Animal Behav. Sci.*, vol. 190, pp. 102–110, May 2017.
[31] J. Abdai, B. Korcsok, P. Korondi, and Á. Miklósi, "Methodological challenges of the use of robots in ethological research," *Animal Behav. Cognition*, vol. 5, no. 4, pp. 326–340, Nov. 2018.
[32] *The Mathis Laboratory of Adaptive Motor Control Deeplabcut*. Accessed: Sep. 9, 2021. [Online]. Available: http://www.mackenziemathislab.org/deeplabcut
[33] R. Labuguen, D. K. Bardeloza, S. B. Negrete, J. Matsumoto, K. Inoue, and T. Shibata, "Primate markerless pose estimation and movement analysis using DeepLabCut," in *Proc. Joint 8th Int. Conf. Informat., Electron. Vis. (ICIEV) 3rd Int. Conf. Imag., Vis. Pattern Recognit. (icIVPR)*, May 2019, pp. 297–300.
[34] S. Fujimori, T. Ishikawa, and H. Watanabe, "Animal behavior classification using DeepLabCut," in *Proc. IEEE 9th Global Conf. Consum. Electron. (GCCE)*, Oct. 2020, pp. 254–257.
[35] P. Ekman and W. Friesen, *Facial Action Coding System: A Technique for the Measurement of Facial Movement*. Palo Alto, CA, USA: Consulting Psychologist Press, 1978.
[36] B. M. Waller, E. Julle-Daniere, and J. Micheletta, "Measuring the evolution of facial 'expression' using multi-species FACS," *Neurosci. Biobehavioral Rev.*, vol. 113, pp. 1–11, Jun. 2020.
[37] P. H. Andersen, S. Broomé, M. Rashid, J. Lundblad, K. Ask, Z. Li, E. Hernlund, M. Rhodin, and H. Kjellström, "Towards machine recognition of facial expressions of pain in horses," *Animals*, vol. 11, no. 6, p. 1643, Jun. 2021.
[38] D. Vincze, M. Gacsi, S. Kovacs, P. Korondi, A. Miklosi, and M. Niitsuma, "Towards the automatic observation and evaluation of ethologically inspired human-robot interaction," in *Proc. IEEE/ASME Int. Conf. Adv. Intell. Mechatronics (AIM)*, Jul. 2021, pp. 586–591.
[39] M. Dziergwa, M. Kaczmarek, P. Kaczmarek, J. Kędzierski, and K. Wadas-Szydłowska, "Long-term cohabitation with a social robot: A case study of the influence of human attachment patterns," *Int. J. Social Robot.*, vol. 10, no. 1, pp. 163–176, Jan. 2018.
[40] M. M. A. de Graaf and S. B. Allouch, "Exploring influencing variables for the acceptance of social robots," *Robot. Auto. Syst.*, vol. 61, no. 12, pp. 1476–1486, Dec. 2013.
[41] M. M. A. de Graaf, S. Ben Allouch, and J. A. G. M. van Dijk, "Why would i use this in my home? A model of domestic social robot acceptance," *Hum.–Comput. Interact.*, vol. 34, no. 2, pp. 115–173, Mar. 2019.
[42] M. K. Strait, C. Aguillon, V. Contreras, and N. Garcia, "The public's perception of humanlike robots: Online social commentary reflects an appearance-based uncanny valley, a general fear of a 'technology takeover' and the unabashed sexualization of female-gendered robots," in *Proc. 26th IEEE Int. Symp. Robot Human Interact. Commun. (RO-MAN)*, Sep. 2017, pp. 1418–1423.
[43] J. P. Hanby, M. J. A. Simpson, B. J. T. Morgan, and J. Hall-Craggs, "Visualizing interaction and sequential data in animal behaviour: Theory and application of cluster-analysis methods," *Behaviour*, vol. 56, nos. 1–2, pp. 1–43, 1976.
[44] K. R. Finn, M. J. Silk, M. A. Porter, and N. Pinter-Wollman, "The use of multilayer network analysis in animal behaviour," *Animal Behav.*, vol. 149, pp. 7–22, Mar. 2019.
[45] M. D. S. Ainsworth and B. A. Witting, "Attachment and exploratory behavior of one-year-olds in a strange situation," in *Determinants of Infant Behaviour*, vol. 4, B. M. Foss, Ed. London, U.K.: Methuen, 1969, pp. 113–136.
[46] B. Nagy and Š. Gubo, "Intelligent space environment for ethorobotics," *Recent Innov. Mechatronics*, vol. 5, no. 1., pp. 1–5, Jul. 2018.
[47] *Optitrack Motive*. Accessed: Oct. 5, 2021. [Online]. Available: https://optitrack.com/software/motive/
[48] Y. Sun, J. Xu, G. Lin, W. Ji, and L. Wang, "RBF neural network-based supervisor control for maglev vehicles on an elastic track with network time delay," *IEEE Trans. Ind. Informat.*, vol. 18, no. 1, pp. 509–519, Jan. 2022.
[49] Y. Sun, J. Xu, H. Wu, G. Lin, and S. Mumtaz, "Deep learning based semi-supervised control for vertical security of maglev vehicle with guaranteed bounded airgap," *IEEE Trans. Intell. Transp. Syst.*, vol. 22, no. 7, pp. 4431–4442, Jul. 2021.
[50] Z. Xu, P. Yang, K. Hu, B. Xu, and H. Li, "Deep learning control model for adaptive optics systems," *Appl. Opt.*, vol. 58, no. 8, pp. 1998–2009, Mar. 2019.
[51] X. Liu, Z. Liu, G. Wang, Z. Cai, and H. Zhang, "Ensemble transfer learning algorithm," *IEEE Access*, vol. 6, pp. 2389–2396, 2017.
[52] L. Prechelt, "Automatic early stopping using cross validation: Quantifying the criteria," *Neural Netw.*, vol. 11, no. 4, pp. 761–767, 1998.
[53] *Solomon Coder*. Accessed: Sep. 15, 2021. [Online]. Available: https://solomon.andraspeter.com/



BALÁZS NAGY was born in Veszprém, Hungary, in 1991. He received the B.S. and M.S. degrees in mechatronics engineering from the Budapest University of Technology and Economics, Budapest, where he is currently pursuing the Ph.D. degree in mechanical engineering.

Since 2015, he has been working with the Department of Mechatronics, Optics, and Mechanical Engineering Informatics, Budapest University of Technology and Economics. His research interests include ethorobotics and deep learning applications in ethology and sensor signal processing.

Mr. Nagy's awards and honors include the Suzuki Fellowship, the Fellowship granted by the Republic of Hungary, and the Fellowship of the New National Excellence Program provided by the Hungarian Ministry of Human Capacities.

PÉTER KORONDI (Senior Member, IEEE) received the Dipl.Eng. and Ph.D. degrees in electrical engineering from the Technical University of Budapest, Budapest, Hungary, in 1984 and 1995, respectively. Since 1986, he has been with the Budapest University of Technology and Economics (formerly the Technical University of Budapest). From April 1993 to April 1995, he worked with the laboratory of Prof. Harashima and Prof. Hashimoto at the Institute of Industrial Science, The University of Tokyo, Japan, where he continues to spend a month each year, working on joint research. As a result of this cooperation, the Intelligent Integrated System Japanese–Hungarian Joint Laboratory was founded in 2001. His research interests include cognitive telemanipulation and ethorobotics. He is a Founding Member of the International PEMC Council.