Development of a new 3D tracking system for multiple marmosets under free-moving conditions

NEW TRACKING SYSTEM FOR FREE-RUNNING CONDITIONS OF MARMOSETS

Terumi Yurimoto[1], Wakako Kumita[1], Kenya Sato[1], Rika Kikuchi[1], Yusuke Shibuki [1], Rino Hashimoto[1], Michiko Kamioka[1], Yumi Hayasegawa[1], Eiko Yamazaki[1], Yoko Kurotaki[2], Norio Goda[3], Junichi Kitakami[4], Tatsuya Fujita[5], Takashi Inoue[1], Erika Sasaki[1*]

[1]Department of Marmoset Biology and Medicine, Central Institute for Experimental Animals, Kawasaki, 210-0821 JAPAN

[2]Center of Basic Technology in Marmoset, Central Institute for Experimental Animals, Kawasaki, 210-0821 JAPAN

[3]Public Digital Transformation Department, Hitachi, Ltd., Shinagawa, 140-8512 JAPAN

[4]Vision AI Solution Design Department Hitachi Solutions Technology, Ltd., Tachikawa, 190-0014 JAPAN

[5]Engineering Department Eastern Japan division, Totec Amenity Limited, Shinjuku, 163-0417 JAPAN

*Corresponding author: Erika Sasaki (esasaki@ciea.or.jp)

Abstract

Assessment of social interactions and behavioral changes in nonhuman primates is useful for understanding brain function changes during life events and pathogenesis of neurological diseases. The common marmoset (*Callithrix jacchus*), that lives in a nuclear family like humans, is a useful model, but long-term automated behavioral observation of multiple animals has not been achieved. Here, we developed a Full Monitoring and Animal Identification (FulMAI) system for long-term detection of three-dimensional (3D) trajectories of each individual in multiple marmosets under free-moving conditions by combining video tracking, Light Detection And Ranging, and deep learning. Using this system, identification of each animal was more than 97% accurate. Location preferences and inter-individual distance could be calculated, and deep learning could detect grooming behavior. The FulMAI system allows us to analyze the natural behavior of individuals in a family over their lifetime and understand how behavior changes due to life events together with other data.

Introduction

Behavioral analyses are important for understanding brain function changes during development and aging and for evaluating disease onset in model animals. In particular, nonhuman primates are often used to study brain function, and behavioral characteristics provide an indispensable source of data for hypothesis testing[1]. Nonhuman primates share anatomical and physiological features with humans, especially in specialized functions acquired through unique cerebral expansions, such as tool use, language (for example, vocal communication), and self-awareness[2, 3, 4, 5]. The common marmoset (*Callithrix jacchus*) is a small nonhuman primate with behavioral and social characteristics resembling those of humans. They are diurnal, form monogamous families, engage in altruistic behaviors such as food sharing, and cooperate with all family members to raise their youngest offspring[3, 6, 7]. As marmoset body size and social units are small, maintaining complete social units for research in a laboratory is easier than that for other primates. Therefore, marmosets are an ideal model for studying social behaviors in social units[3]. Recent developments in genetic modification techniques in marmosets have enabled the production of various disease-model marmosets, such as those for neurodegenerative diseases[8, 9, 10]. Using marmosets as model animals has advantages over rodents and may help elucidate basic gene functions in the brain, clarify molecular mechanisms underlying neurological disease onset using genetic modifications, and aid in implementing cognitive neuroscience techniques in primate research[5, 11]. For these purposes, national brain research projects

using marmosets have been conducted in the US and Japan[12, 13].

To evaluate long-term brain function changes due to development, aging, and disease onset using marmosets as a model, it is important to capture changes in behavior over a lifespan, and an analysis of multiple marmosets under free-moving conditions is necessary to capture changes in social behavior. Therefore, a novel system is required to detect and quantify the individual behaviors of multiple marmosets under free-moving conditions in real time, both automatically and accurately. Automated behavior analysis in a home cage enable observations of more natural animal behavior[14, 15, 16]. This analysis has the advantages of being able to ignore environmental condition, monitor social behavior in multiple animals, and analyze changes in social behavior over a relatively long period of time[17].

The various systems used for analyzing animal behaviors under free-moving activities include systems suitable for three dimensions (3D) tracking and behavior classification (DANNCE[18], FreiPose [19], and MarmoDetecotor[20]), for generating pose estimation (DeepLabCut, SLEAP[21], MARS[22], DANNCE, FreiPose, and OpenMonkeyStudio[1]), for analyzing behavior in home cages (DeepLabCut [23, 24], B-SoiD [17], SLEAP, and MARS), for analyzing multiple animals (DeepLabCut, MiceProfier[25], SLEAP, and OpenMonkeyStudio), for finding novel behavioral abnormalities by unsupervised clustering (MoSeq[26], B-SoiD, and FreiPose[19]), and for use across animal species (DeepLabCut, DeepBhvTracking[27], B-SoiD, and DANNCE). These systems have been generally developed for mice;

68 only DeepLabCut, DeepBhvTracking, DANNCE, and MarmoDetector[20] have been adapted to
69 marmosets, which move relatively quickly in 3D. A previous study used the pose estimation system
70 to analyze the behavior of two marmosets for nine hours using individual identification markers[23].
71 These analysis systems have observed only a small proportion of an animal's life and longitudinal
72 observation over each marmoset's lifetime (15 years) or more than months of observing a family have
73 not been achieved. This is possibly because automatically detecting marmoset trajectories, behaviors,
74 and interactions between individuals under free-moving conditions in multiple animals is especially
75 challenging. For example, marmosets are often close to each other when grooming, mating, or
76 fighting[28]. This proximity makes it difficult to detect individual bodies, which complicates individual
77 tracking and behavior detection.

78 In this study, we developed a behavioral analysis system named FulMAI (Full Monitoring
79 and Animal Identification), that can record and analyze behavior of an animal living among multiple
80 animals over their lifetime. The concept of this system is to detect changes in behavior over an
81 individual's lifespan and understand the relation of changes in brain function. FulMAI is a 3D tracking
82 system that uses cameras, light detection and ranging (Lidar) devices, and deep learning to
83 simultaneously track a marmoset family of three individuals in one cage. Furthermore, to take
84 advantage of stress-free behavior analysis in the home cage, we also developed a system for individual
85 identification via facial recognition using deep learning. The FulMAI system can analyze the location

and the period each animal stayed in the cage and the distance between individuals. Furthermore, the time spent in social behavior is an index for analyzing animal interactions. Therefore, as an example of social behavior detection, we developed automatic grooming behavior detection system using deep learning. The technology developed in this study would be applicable in detecting important behavioral changes caused by development, aging, and various diseases in small, nonhuman model primates.

Results

Marmoset facial recognition by deep learning

Five marmosets of the same generation (1–2 years old) were used to examine whether deep learning for face identification could be adapted to marmosets. These marmosets were kept in different cages in a conventional breeding colony and facial images were collected individually. To evaluate the trained models, 500 new facial images (100 for each animal) were obtained from the videos not used for training. The evaluation metrics for deep learning were as follows: specificity of 99.8%–100%, accuracy of 99.4%–99.8%, recall of 98.0%–100%, precision of 99%–100%, and an F-measure of 98.5%–99.5% (Table 1). The overall identification accuracy (total number of correct images/total number of images) of the five marmosets was 99.2%.

Subsequently, face identification accuracy was examined in a family of three, that was

raised in the set-up home cage. The trained model was evaluated using data that were not used for training. For the evaluation, 5,600 facial images (1,400 for each class), three classes including each marmoset, and an unknown class were used. The unknown class was used to exclude inappropriate images, such as unclear images, images of the back of the head and faces in profile, and images not showing faces. The evaluation metrics for deep learning for individual identification were as follows: specificity of 98.7%–99.7%, accuracy of 98.6%–99.2%, recall of 96.1%–98.8%, precision of 98.4%–99.1%, and an F-measure of 97.3%–98.3% (Table 2). The overall percentage of correct identifications was 98.0%.

Multiple animals' 3D tracking

To optimize the Lidar for laser measurement and video recording, a part of the cage made of metal mesh was replaced with acrylic panels, following which suitable measurement results were obtained (Fig. 1a). The laser beam distance needed to be sufficiently narrow and cover the entire cage; thus, the Lidars were placed 1000 mm from the front of the cage (Fig. 1b). All 3D tracking and grooming behavior detection tests were conducted using data from one family consisting of three marmosets (I5072M/father (I5072M/f), I5894F/mother (I5894F/m), and I940M/juvenile (I940M/j)) in a cage.

The 3D coordinate information of marmosets was obtained using video tracking and Lidar (Fig. 2a-f, and Mov. S1). While the spatial resolution of Lidar was not enough to detect each marmoset and its behavior, it was sufficient to detect the body and tail. Therefore, the centroids of the point

clouds were used as animal locations for 3D trajectory.

A raw image taken by the upper left video camera captured the entire cage and the three marmosets (Fig. 2a), and the data was acquired by Lidar simultaneously (Fig. 2b). The centroid of cluster by Lidar was indicated in white points (Fig. 2b), and each point was given a sequential tracking ID number during background processing. A new number was assigned each time when the trajectory was broken. The 2D pixel locations of the marmosets were then detected using an object detection algorithm (Yolov3[29]) in the video image (Fig. 2c). The same process was applied to the four cameras to determine the locations of the marmosets for each camera. The 3D location of each marmoset was calculated using pre-calibrated camera positions and the pixel coordinates of the marmosets (Fig. 2d). The extracted face images of marmosets by Yolo were identified using the trained VGG19 model, and individual IDs were assigned to each marmoset by facial recognition (Fig. 2e, f). These series of processed video data are shown in Mov. S1. In some cases, the marmosets could not be recognized by video tracking when they moved too fast or behind an object. In such cases, the trajectory was interpolated by Lidar (Mov. S2, S3). Fig. 3 shows 5 min of tracking trajectories without face identifications (Fig. 3a, b) and with face recognition (Fig. 3c). When Lidar alone or Lidar and video tracking were performed, the tracking serial numbers changed frequently, and it was impossible to continuously track the individual (Fig. 3a, b). However, by combining face recognition with Lidar and video tracking, the trajectory of each animal was continuously detected, and marmoset IDs were

assigned (Fig. 3c).

Trajectory analysis using the 3D tracking system

Using the results of 3D trajectories from this 3D tracking system, we analyzed where and for how long each animal stayed in the cage during the one-hour period of the analyses (Fig. 4). This data indicated that I5072M/f stayed on the lower floor of the cage for 88.8% (23642/26626 frames) of the one hour, whereas both I5894F/m and I940M/j stayed on the lower floor for less than 0.5% (128/25597 and 12/24234 frames) of the one hour. In contrast, they stayed on middle floor for half of the one hour and actively moved in various areas in the cage except the lower floor rest of the time (Fig. 4a-c, Table 3, Supplementary Fig. 1). Furthermore, the individual distances were analyzed using the location data (Fig. 4d). During this time, the distance between I940M/j and I5894F/m was less than 0.5 m 25.9% (5044/19501 frames) of the total time. However, a distance of less than 0.5 m between I940M/j-I5072M/f and I5072M/f-I5894F/m was detected 1.0% (203/20409 frames) and 2.2% (475/21599 frames) of the time, respectively. Conversely, the distance between I940M/j-I5894F/m was more than 1 m 2.0% (391/19501 frames) of the time, whereas I940M/j-I5072M/f and I5072M/f-I5894F/m were separated by more than 1 m distance 20.0% (4081/20409 frames) and 11.0% (2364/21599 frames) of the time, respectively. In this family, I940M/j and I5894F/m spent more time together than I5072M/f.

The same analyses were performed for long-term analysis on a different day during the daytime, 7:00 am–7:00 pm. After 6:00 pm, all marmosets moved to the bed and clustered together and

158 their faces were not observed; therefore, the data were excluded from the analysis. The detailed
159 analyses indicated that, among the family members, I940M/j stayed on the lower floor of the cage the
160 longest, unlike the results of the 1-hour analysis (Supplementary Fig. 2). The time spent on the lower
161 floor of the cage between 7:00 am and 6:00 pm for each individual was 8.6% for I5072M/f, 13.0% for
162 I5894F/m, and 23.5% for I940M/j. Throughout the day, I940M/j spent the most time on the middle
163 floor (27.7%), I5072M/f on the upper floor (29.7%), and I5894F/m on the upper floor (29.9%).
164 Furthermore, locations of each animal were analyzed hourly for 12 hours (Supplementary Figs. 2, 3).
165 I5072M/f also spent most of his time in the upper area in the morning (9:00 am; 54%), but gradually
166 changed his activity area to the middle area in the afternoon (1:00 pm; 51%) and the lower area in the
167 evening (4:00 pm; 45%). I5894F/m rarely stayed in the middle area but moved to the upper section in
168 the morning (9:00 am; 48%) and to the lower section in the afternoon (3:00 pm; 55%). I940M/j moved
169 most evenly throughout the day among the three in the cage. At night, all individuals moved to the
170 upper bunk to their beds (Supplementary Fig. 2).

171 The distance between the marmoset pairs was also analyzed throughout the day. The
172 results showed that I5072M/f-I5894F/m were closer than 0.5 m for 21.1% of the daytime. However,
173 I5072M/f-I940M/j and I5894F/m-I940M/j were closer than 0.5 m for 19.6% and 16.7%, respectively.
174 Additionally, the closest and farthest pairs per hour switched frequently throughout the day
175 (Supplementary Fig. 5).

Detection of grooming behavior

In this system, behaviors could also be detected using video data from the two upper cameras in front of the cage (Supplemental Fig. 6a, b). Grooming behavior, one of the representative social behaviors of marmosets, excluding self-grooming behavior was detected. Training data on grooming behavior were obtained through observations of video data by skilled animal technicians. To increase the accuracy of grooming behavior detection, situations in which marmosets were close to each other but not grooming were rigorously classified by skilled animal technicians as non-grooming images and eliminated from the supervised images. The model was trained by annotated data of the grooming positions in each video frame. Excluding annotation time, 880 hours were spent on training. Creation of a csv file after processing of four movie files took 8 min. Face identification took 20 msec per process, and matching face identification to trajectory took 1 min per 10-min video.

The coordinates of the grooming location and individuals were detected by separately using the 3D tracking system, and determining which individual performed the grooming. Grooming behaviors were observed more often in the early morning (7:30 am–8:00 am), midday (12:30 pm–1:30 pm), and evening (3:30 pm–4:00 pm, 5:00 pm), and all marmosets performed grooming behaviors equally (Fig. 5a). All marmosets groomed at fixed locations (Fig. 5b), namely the two upper areas of the cage. These two locations were different from the sleeping area (Supplemental Fig. 3).

Accuracy verification of the 3D tracking system

To determine the accuracy of the 3D tracking system, marmosets were fitted with colored belly band to identify each animal, and individual tracking was conducted for 60 min. The Mov. S4 shows a portion of this experiment and confirms that the system could track the marmosets accurately even when they crossed. Although the marmoset cage had a swinging hammock and balls as enrichment (Mov. S5), these movements did not affect the tracking accuracy of the marmosets when using a combination of video tracking and Lidar. During this experiment, the marmoset IDs switched to different IDs nine times for more than 5 sec, but seven of the nine ID switching events were resolved within 15 sec. The maximum time length for accurate individual tracking was 19 min. Incorrect marmoset IDs of longer than 5 sec were exhibited over 4.6% of the total 60 min.

The accuracy of the 3D trajectory data analysis in detecting the positions of animals in cages was confirmed by comparing it with visual observation data (Table 3, Supplemental Fig. 2). For example, in the visual observations of an hour, the position of the animals was on the upper wooden bed 9.4% of the time, on the middle floor 63.8%, on the lower wooden bed 25.1%, on the lower floor 0%, and in the other areas 1.7% of the time. Similarly, the system-analyzed data showed that the position of the animals on the upper wooden bed was 16.5%, middle floor was 52.6%, lower wooden bed was 29.1%, lower floor was 0.1%, and other areas was 1.7%. A strong significant correlation was found between the visual observation and system data ($R^2 = 0.988$, $P < 0.01$; Fig. 6a).

The accuracy of grooming behavior detection by the machine was compared with that of

visual observation for one hour (Fig. 6b). Each 30-sec video was considered as one unit. Twenty-four grooming units were visually detected, and 21 units were detected by the machine. Among these, 19 units were detected by visual observation and machine (recall: 79.1%). In the system, 19 of 21 units were confirmed as actual grooming behaviors (precision: 90.5%). The five missed units were difficult to detect because of the placement of animals behind an object: four had marmosets grooming behind the feeder, and one had marmosets grooming in the blind spot of the bed. Conversely, two units were detected by the machine but not by direct observation. During these events, the marmoset nuzzled its face into the other's tail or placed its hand on the other's shoulder. These actions seemed to be part of grooming behavior; however, observers did not judge them as grooming behaviors because they were not typical grooming actions. In one hour, there were 10 units during which the distances between marmosets were close, and false positives were detected in only one unit, with a duration of 30 sec (recall: 90%, Fig. 6b). The system also worked well in higher interaction situations, such as grooming. A grooming situation is shown in Mov. S6. The two animals were very close to each other and the animal ID switched frequently, but both were involved in grooming. After grooming, accurate individual identification was resumed as soon as the face was visible, as shown in Mov. S4.

Discussion

In this study, we constructed the FulMAI system, a new automatic 3D tracking system for multiple

marmosets under free-moving conditions, using a novel approach that combines Lidar, video tracking, and deep learning face recognition. These are all existing technologies, but this is the first study combining them for 3D tracking of multiple unlabeled nonhuman primates with relatively rapid movements in free-moving conditions. The FulMAI system allowed us to analyze the natural behavior of individuals in a family for a month. Lidar and Yolo, the object detection algorithm, were fast programs that could be run constantly for a month with little delay. In addition, Yolo's light memory usage and easy integration with other systems will allow us to obtain more long-term and diverse data through well-planned operations in the future. Based on these findings, FulMAI is expected to record the lifetime behavior of marmosets with a lifespan of approximately 15 years and match it with MRI and other data to determine what changes occurred in the brain during life events and transformed their behavior.

In preliminary studies, we used metal mesh doors, but these reflected the laser, and the marmosets' location could not be detected. The cage doors were changed to acrylic panels to make them suitable for the analysis (Fig. 1). The cages can be divided into four compartments. As shown in Supplemental Fig. 1, the system accurately tracked each animal even though the cage was divided. Therefore, the cage size can be changed to the extent that it fits within the camera's view. In contrast, thick poles should be avoided in front of the cage because they increase blind space for both the camera and the Lidar.

Although either laser or video data can be used as the primary data for this system, we used video tracking as a base for detecting marmoset positions and then completed the missing information using Lidar, thereby compensating for the lost marmoset trajectory detections by video and false recognitions by Lidar. Owing to its detection speed and resolution, Lidar is suitable for application to fast-moving objects. However, it falsely recognized objects that were similar in size to a marmoset, such as a marmoset tail, feeder boxes, or colored balls. On the other hand, a hammock in the family cage, which was included as an environmental enrichment and was not firmly fixed to the cage, was not mis-detected by either Lidar or video because of the differences in size and shape from those of the marmoset, and it did not affect the tracking of the marmosets. Video tracking does not have this issue because it uses color image data to identify objects (Mov. S5). However, video tracking lost detection when marmoset motion was quicker than the camera's shutter speed or when the marmosets were behind obstacles (Mov. S2, S3). By compensating for the respective shortcomings of Lidar and video tracking and using each to complement the trajectories of the other, a tracking system for small, high-speed objects in three dimensions was realized. In addition, because both Lidar and video tracking can be processed at high speed without a time lag, data is not accumulated for subsequent analysis, which is favorable for long-term analysis. Indeed, the 3D trajectories and grooming behaviors of three marmosets could be continuously analyzed in their home cage for over six months without data accumulation. The animals were healthy throughout the experiment, proving the safety of our

system. A similar 3D tracking system using Kinect senser or cameras has been reported for various animal species, including marmosets[1, 20, 23, 24, 30, 31]. However, none of the existing 3D tracking systems have achieved such a long-term analysis, whereas FulMAI is the only behavioral analysis system for longitudinal observation over each marmoset's lifetime in the family. One of the advantages of FulMAI is the flexibility in terms of cage size and structure as long as acrylic panels are located in front of the camera and Lidar, and it can be applied to a wide variety of animal species. This system would also be useful for other animal species that move as quickly as marmosets in 3D, such as tamarins, free-swimming fish, and quick-flying birds and bats.

Our system successfully tracked multiple marker-less marmosets for the first time, which is another advantage of our system over existing systems. The color marking causes animal stress because it requires capture or anesthesia, and periodic hair dyeing requires time and effort, which is unsuitable for analyzing marmosets' behaviors in the long-term. The face identification of marmosets, which we were able to achieve for the first time, can also be undertaken independently and can be applied to other motion analysis systems, such as DeepBhvTracking[27].

When marmosets are kept as a family, they repeat the behavior of huddling and then suddenly dispersing several times a day. By this characteristic behavior, individual traceability is lost from the trajectory. Therefore, we combined our system with facial recognition, using deep learning to detect and track each animal in a family group, since the most distinctive feature of the marmoset's

284 body is its face. Usually marmoset individual identification is performed using microchips, hair
285 clippings, hair dyes, or necklaces[32]. Microchips and necklaces are unsuitable for MRI analysis, hair
286 clippings and hair dyes require periodic maintenance, and necklaces pose the risk of tightening around
287 the neck as they grow or of getting caught. In one part of this study, the marmosets were fitted with
288 colored belly band for visual tracking, but I5894F/m seemed uncomfortable with the band and was
289 constantly fidgeting with it. Thus, artificial markers have the risk of causing unexpected behavioral
290 changes. Therefore, one of the goals of this study was to achieve marker-less identification of
291 individuals.

292 The marmoset family used in this study had no blood relationships with other family
293 members since the juveniles (I940M) were obtained by an embryo from another embryo donor pair
294 and was transferred into the female's uterus (I5894F; foster mother)[10, 33]. Regarding facial
295 identification, the most accurate results within the family were those for I940M/j (Table 2). Although
296 I940M/j was beyond sexual maturity, younger animals have easily identifiable distinctive facial
297 features. In addition, when using five animals of the same age (1–2 years old), the animals could be
298 individually identified; therefore, highly accurate classification is possible by facial recognition even
299 when the animals are of similar age (Table 1). Incorporating facial recognition results into the
300 trajectory enabled long-term tracking, even for multiple animals in the same cage (Fig. 3). We
301 examined the accuracy of marmoset tracking using this system by color labeling the marmosets and

comparing this to visual tracking (Mov. S4, Table 3). This experiment confirmed that it is possible to track animals using facial recognition with 95% accuracy, without labeling marmosets or relying on color information.

In addition, using the 3D coordinates data, this system could calculate where each animal stayed in the cage, for how long it stayed (Fig. 4a-c, Table 3, Supplemental Figs. 2, 3, 4), and the distance between individuals (Fig. 4d, Supplemental Fig. 5), which suggests interactions between individuals. Although switching of the marmoset IDs was observed several times during the hour-long experiment, results of visual observations and system analysis for where each marmoset existed in terms of 3D coordinates in the cage, and how long each marmoset spent in that area were highly correlated ($R^2 = 0.988$, Fig. 6a, Supplemental Fig. 2). This indicated that the slight ID switching did not affect the experimental results, and our tracking system was sufficiently accurate for automatically analyzing marmosets' 3D coordinates in the cage.

Marmosets often change their preferred position and the individuals that they spend time with during the day. For example, the male marmoset (I5072M/f) preferred the upper area in the morning, but in the afternoon, it preferred the lower bed. This result was not apparent in the 1-hour analysis compared with observations. Furthermore, one-hour and 12-hour analyses of position in the cage and distance between marmoset individuals showed different results. For example, a one-hour analysis showed that I5072M/f preferred the cage's lower floor and remained there 80% of the analysis

320 time (Table 3). However, a 12-hour analysis indicated that I940M/j spent the longest time (23.5%) on
321 the lower floor (Supplemental Fig. 3, 4). The analyses of the distance between individuals showed that
322 I5894F/m-I940M/j were within 0.5 m for 25.9% of the one-hour analysis, but the 12-hour analysis
323 indicated that the I5072M/f-I5894F/m remained closer than I5894F/m-I940M/j or I5072M/f-
324 I5894F/m. These results suggest that the results based on human observation for a limited time and
325 those based on machine observation throughout the day may differ. Indeed, when we directly observed
326 this family, I5072M/f and I5894F/m were usually far apart, but a 12-hour analysis suggests that the
327 I5072M/f and I5894F/m spent time close together. Therefore, it should also be considered that people's
328 presence alters marmoset behavior. To accurately understand marmoset behavior, it is necessary for
329 the behavior analysis system to constantly analyze a marmoset's life with the same evaluation criteria.
330 Our 3D tracking system can continuously analyze animal positions and distances between individuals.
331 In addition, some marmoset facilities, to provide more space or different environments
332 for marmosets, occasionally connect adjacent marmosets' cages[34]. In this study, an extra cage was
333 connected to the home cage. When a marmoset went off-screen due to movement to the connected
334 cage, the tracking of the marmoset was temporarily interrupted. However, other marmosets continued
335 to be tracked, and no other individual was tracked instead of the missing marmoset (Mov. S7).
336 Tracking was resumed when the marmoset returned to the cage.
337 Grooming, a typical social behavior, was detected through deep learning using video

338 analysis. The detector was Yolo, the same one used for the 3D tracking. Comparing detection of the
339 grooming behavior using visual observation by humans and the system showed that grooming
340 behaviors were detected with high precision (90.5%) and recall (79.1%; Fig. 6b). In some cases, the
341 system identified behaviors just before or after the grooming as grooming behaviors. The deep learning
342 system may have detected certain behaviors that might be unique to periods before and after grooming.
343 However, in other cases, grooming behavior could not be detected by the deep learning system because
344 the groomers were located in the camera's blind spot. Therefore, the training dataset must be further
345 updated to detect behavior behind obstacles. Combining the 3D tracking with the grooming behavior
346 detection by deep learning, this system could reveal when and where each marmoset displayed
347 grooming behavior using accurate quantification, and the social relationships among the animals (Fig.
348 5 a, b). The 12-hour grooming behavior analysis in this study indicated that grooming behavior was
349 found to occur at specific locations: upper beds and floor, and at specific times: early morning, midday,
350 and evening. This finding was obtained exclusively using automatic behavior analysis. While there
351 are existing methods for automatically detecting the behavior of a single animal[35], for the first time,
352 our system was able to directly detect the behavior of multiple animals in captivity and identify the
353 individuals and their locations. We aim to identify a variety of behaviors for analyzing lifelong
354 behavior changes. To find unpredictable unknown behavior, clustering by unsupervised learning
355 would be necessary. For unsupervised learning, it is necessary to analyze the videos of each individual

or to create a skeletal model of each marmoset and analyze the movement of each key point. Our study is expected to further develop in the future because of its ability to detect accurate 3D location linked to each individual. For more flexible behavior extraction, this system was made of Robot operating system (ROS) and is extensible to incorporate other Python-based projects, such as DeepLabCut, DeepEthogram[36], and Tweetynet[37]. Because all animal activities are recorded in the FulMAI, each animal behavior in this home cage can be analyzed retrospectively in detail. By comparing with other data, such as imaging, blood biochemistry, or biomarker data, it is also possible to analyze the behavior of model animals from multiple perspectives without overlooking changes in behavior associated with the life events of brain function changes or the onset of disease.

Moreover, Lidar can track the animal locations day and night; therefore, their nocturnal behavior can be observed to obtain primary data. Our preliminary data showed that the combination with a Lidar and starlight camera can capture details of nighttime behavior in a dark room (Supplemental Mov. S8). Sleep disorders have recently been suggested as predictors of Alzheimer's disease and Parkinson's disease[38, 39]. This system would be able to accurately quantify the behavior of marmosets throughout their lives to discover behavioral changes before and after the onset of diseases, even at an extremely early stage.

Recently Calapai et al. reported a new touch screen-based cognitive auditory analysis system that allows marmosets to perform tasks in a free moving environment in a paired home cage[40].

By incorporating such a freely operable cognitive function task device into our system, it is possible to provide a behavior analysis system that comprehensively and continuously analyzes the marmoset's natural behavior, activity, and cognitive functions without inducing stress.

In conclusion, the combination of Lidar, video tracking, and facial recognition enabled successful simultaneous tracking of each marmoset in a family without artificial labeling. Using data from the 3D tracking system we developed, we could measure 3D coordinates, movement trajectories, the quantity of activity, position in the cage, when, and for how long each animal stayed, distances between individuals, and time spent in grooming behavior. These measurements will allow automated analysis of activity and behavior under free-moving conditions throughout the marmoset's lifetime, enabling quantitative analysis of "when, where, and what" of each animal and capturing changes in brain function, including development, growth, aging, and disease.

Materials and Methods

Hardware Setup

The home cage was made of acrylic panels on the front and back, and all food dishes, water jugs, and bedding inside the cage were fixed to a wire mesh on the side to prevent them from moving. This cage was designed to conduct behavioral analysis in a marmoset's home cage (Fig. 1a). Four poles were set up in front of the cage at 1000 mm, and four Lidars (VLP-16, Velodyne Lidar, San Jose, CA, USA)

392 and four cameras (C920, Logitech, Lausanne, Switzerland) were fixed to the poles (Fig. 1b). The Lidar
393 was controlled by motion measurement software, the path traced, and velocity of the marmosets
394 recorded (Hitachi, Ltd, Tokyo, Japan.). The lower Lidar was set at 0.9 m height and the upper Lidar
395 at 1.9 m height. The distance between the left and right Lidar devices was 0.3 m.

396 The laser beam of the Lidar was irradiated radially and adjusted to cover the entire interior
397 area of the cage. The recording parameters of the cameras were 1080p and 60 Hz, and all cameras
398 were adjusted to show the entire cage (Supplemental Fig. 6). The positions of the cage, Lidar, and
399 cameras were calibrated when the cage was moved, and the interiors were changed. Lidar calibration
400 was used to acquire background information, including the cage interior without marmosets.
401 Calibration of the camera was performed by detecting alco-markers at the four corners of the cage.
402 The specifications of the PC used to operate and record the camera and Lidar were corei7, RTX1060,
403 32GB. PC specifications used for analysis were corei7, RTX3090, 64 GB.

404 **Ethical approval**

405 This study was conducted at the animal facility of the Central Institute for Experimental Animals
406 (CIEA), Kawasaki, Japan. The animal experimental protocol was reviewed by the Institutional Animal
407 Care and Use Committee and approved (approval no. 17051) according to the Regulations for Animal
408 Experiments in CIEA, based on the Guidelines for the Proper Conduct of Animal Experiments by the
409 Science Council of Japan (2006).

Animal facilities

The marmosets were kept in a family cage with dimensions of 820 × 610 × 1578 mm. Although there was only one marmoset family in the room, they were able to vocally communicate with the marmosets in other rooms. All cages were equipped with a bunk, and the animal rooms were adjusted to 26–28°C temperature, and 40–60% humidity with a 12:12 h light/dark cycle. The animals tested negative for *Salmonella* spp., *Shigella* spp., and *Yersinia pseudotuberculosis* on annual fecal examination. The study animals were inspected nine months before the study.

Animals

Each marmoset family was housed in a home cage. The observed family consisted of a male (I5072M/f, ten years old), female (I5894F/m, six years old), and a juvenile (I940M/j, one year old). No artificial or non-artificial physical features were used for individual identification, such as hair dye, necklaces, piercings, or other visible signs, except during 3D tracking of marmosets during the color marker experiments. In conventional animal care, a microchip is implanted subcutaneously for individual identification. The individual I940M/j was born by embryo transfer to the mother's uterus[10], and there was no blood relationship within the third degree of the family.

Five healthy marmosets approximately 1–2 years old, one male and four females, were used in the preliminary study to develop deep learning facial recognition (I774 one year old, I6695 one year old, I6708 one year old, I6677 one year old, and I893 two years old). There were no artificial

or visible physical characteristics, such as hair dye, necklaces, or piercings, and each marmoset was not related to any other marmoset within the third degree of kinship.

System workflow

The overview of this system is shown in Supplemental Fig. 7. The videos and Lidar data were processed on the same PC, and acquisition times were perfectly synchronized for at least one month. Yolo, an object recognition algorithm, detected the marmoset's face, body, actions, and environmental enrichments (ball and hammock) in each video frame. Body coordinates were combined with coordinates using Yolo and Lidar to calculate a short 3D trajectory. Furthermore, the cropped face images were identified as individuals using a convolutional neural network, and individual IDs were applied to the short-term 3D trajectories to achieve long-term 3D tracking of each individual. Behaviors were detected simultaneously with the same model of Yolo, and each animal's behavior was linked to the nearest marmoset ID.

Facial recognition

The cage (W820 × D610 × H1578 mm) was divided into four compartments to isolate each marmoset for obtaining training images by video camera. Each animal was housed in a separate cage, and videos were recorded from the cage front for 3–6 hours. The front door of the cage was then replaced with clear acrylic panels. From these video frames, the face images were cropped only in the face area for the training images, and these data were augmented by rotation. Using 1400 facial images in each

class for training, frontal images of the face were classified into four or five classes using the VGG19 model network. In the preliminary test, the five classes comprised five marmosets. For the family group, the four classes included those members and an unknown class to exclude images inappropriate for classification, and 1000 images were used for training.

Model training was performed by fine-tuning a pre-trained VGG19 model (learning rate 0.001, fixed number of layers 12, BATCH_SIZE 32)[41] (Supplementary Data S2). VGG19 is a network model proposed by the Visual Geometry Group at the University of Oxford in 2014 that obtained accurate classification performance on the ImageNet dataset. The pre-trained models used in this study were trained using the ImageNet dataset.

3D tracking

Lidar and video tracking performed 3D tracking. A Lidar measures the distance to an object by irradiating a laser and analyzing the reflected laser. Marmoset detection using Lidar involves the following steps: background information removal, clustering, and selection of marmoset clusters[42]. Background removal was performed during calibration. The observed points were moving points belonging to objects whose spatial positions changed and static points from the background environment. Marmosets were then selected from among all the moving clusters. The Euclidean cluster extraction algorithm was used for clustering, with a minimum cluster size of S1 and maximum cluster size of S2 (S1: 100 mm, S2: 500 mm)[43]. The centroid of the cluster is regarded as the position

464 of the marmoset for tracking. Three-dimensional video tracking was performed using four time-
465 synchronized video cameras. The positions of the whole body and head of the marmoset were detected
466 in each video frame using Yolo, an object detection algorithm, and the 3D position was calculated
467 from each camera information[29]. The Yolo model was trained using 1164 images in which the whole
468 body and head area were annotated (learning rate: 0.0003, momentum: 0.9, batch: 64). For tracking,
469 the primary information used was the 3D video tracking coordinates, and the information was
470 supplemented with Lidar coordinates when video coordinates were missing. When the primary video
471 information was lost, the system searched for available Lidar coordinates close to the coordinates just
472 before the location was lost and switched to Lidar tracking. When Lidar coordinates were available,
473 continued tracking was undertaken using only Lidar. When the camera re-detected the coordinates,
474 the tracking was switched back to the camera again. The switching was set to occur when the
475 coordinate relationship between the camera and Lidar coordinates was fixed.

476 The program for linkage between video tracking and Lidar for individual information was
477 produced by Hitachi, Ltd. When the marmoset coordinates were detected outside the cage, they were
478 considered as missing values.

479 **Detecting grooming behavior**

480 The software Yolo was used to detect grooming behavior. A total of 252 annotated image files in which
481 grooming marmosets were surrounded by a minimum rectangle were prepared as training images, and

482 the model was trained using the dataset. Grooming images were strictly categorized by skilled animal
483 technicians to exclude situations in which marmosets did not groom. The trained model analyzed the
484 top two camera view images and detected grooming behavior in each video frame. To improve the
485 accuracy by removing noise, a Gaussian-weighted moving average of 50 sec window and 5 sec slide
486 was applied to the frequency of grooming occurrence each time to create the probability distribution
487 of grooming occurrence. The same process was performed for the left and right cameras and the
488 calculated probability distributions were averaged. When the average value exceeded a certain
489 threshold, grooming was considered to have occurred (Supplementary Data S2). To test the accuracy
490 of this detection system, skilled breeding staff visually detected marmoset grooming in one hour of
491 video that was not used for training data. This visual result was used as the true label, and the accuracy
492 and repeatability of grooming detection were calculated in 30 sec increments of time. Because
493 grooming was analyzed with two cameras, it was possible to link location and time. Grooming
494 locations were calculated using the same method as used for video tracking. The grooming time was
495 based on the video file.

496 **Statistical analysis**

497 Pearson's correlation coefficient was calculated using GraphPad Prism version 9 (GraphPad Software,
498 La Jolla, CA, USA).

499

500

Acknowledgments

This research was partially supported by the Brain Mapping by Integrated Neurotechnologies for Disease Studies (Brain/MINDS), "Study of developing neurodegenerative model marmosets and establishment of novel reproductive methodology (JP19dm0207065)" from the Japan Agency for Medical Research and Development (AMED) and Grant-in-Aid for Scientific Research (A) "Development of an Automated Behavioral Analysis System for Evaluation of Genetically Modified Disease Model Marmosets" JSPS KAKENHI Grant Numbers JP21H04756 to ES and Grant-in-Aid for Early-Career Scientists "Effects of transient alcohol exposure to the fetus in early pregnancy on brain development." JSPS KAKENHI Grant Number JP20K16908 to TY.

Author contributions

T.Y. and E.S. conceived and designed the study. T.Y., N.G., J.K., and T.F. performed the setup of the measurement system. R.K., R.H., M.K., Y.H., and E.Y. performed the annotation of the training data. T.Y., Y.S., R.H., J.K., and T.F. processed the data, trained the models, and analyzed the results. T.Y. and R.K. performed the experiment. T.Y., W.K., K.S., N.G., T.I., and E.S. discussed and interpreted the results. T.Y., T.I., and E.S. wrote the manuscript. E.S. supervised the project. All authors commented on the manuscript.

Competing interests

NG is an employee of Hitachi, Ltd. JK is an employee of Hitachi Solutions Technology Ltd. TF is an employee of Totec Amenity Limited. The remaining authors declare no competing interests.

Resources Availability

Information and requests for the program for linkage between video tracking and Lidar should be directed to and will be fulfilled by Norio Goda (norio.goda.ws@hitachi.com). Requests for other resources and information should be addressed to Erika Sasaki (esasaki@cieea.or.jp).

Data and code availability

The readily tested code and data has been presented in the Supplemental information (Supplementary Data S1, S2).

Figure legends

Figure 1. Installation of hardware for the behavior analysis system

(a) A home cage with acrylic panels on the front and back and metal mesh on the sides. (b) Schematic diagram of light and ranging (Lidar) systems and cameras installed in front of the cage. Lidars and cameras were installed 1000 mm in front of the cage.

Figure 2. 3D tracking of marmosets using Lidar and video tracking

(a) Raw video frame. (b) Detected 3D coordinates of marmosets by Lidar. White point: centroid of marmosets; white box: virtual space of the cage. (c) Detected location of marmosets in video frame by Yolo. Rectangle: detected marmosets. (d) Calculated 3D coordinates of marmosets by video tracking, colored point: coordinate of marmoset. (e, f) Individual information added to c and d using

face identification, green: I5072M/f, red: I5894F/m, yellow: I940M/j.

Figure 3. 3D position trajectory of a marmoset over a period of five minutes

(a) Marmoset trajectories analyzed by Lidar. White points indicate centroid of marmosets. (b) Lidar and video tracking. Different colors were assigned to each tracking serial number that numbered the background. (c) Lidar and video trajectory tracking combined with face recognition. Green: I5072M/f; red: I5894F/m; yellow: I940M/j.

Figure 4. Place preference of marmosets over a period of 1 hour

Color map of the time spent by marmosets in the cage for one hour (a-c). Locations where marmosets spent more than 1% of one hour are shown in red, and others are shown as color bars. (a) I5072M/f, (b) I5894F/m, (c) I940M/j. (d) Histogram of the distance between each family member. Blue bars indicate the individual distances between I940M/j and I5072M/f, orange bars indicate the individual distances between I940M/j and I5894F/m, and green bars indicate the individual distances between I5072M/f and I5894F/m.

Figure 5. Detection of grooming behavior for a day

(a) Number of grooming behaviors detected per day. Bars show the number of frames detected each 5 min. (b) The 3D coordinates where the grooming behavior of each individual was detected are shown. Red indicates I940M/j, green indicates I5072M/f, and blue indicates I5894F/m.

Figure 6. Comparison of human visual observations and the system

(a) The time the marmosets stayed in a particular location was compared between visual observations and those by the system. The marmosets stayed in four locations: upper bunk, middle floor, lower bunk, and lower floor. Dotted lines indicate regression lines. (b) Each cell represented 30 s. Blue cells: Visual observation detected grooming at least once in 30 s. Yellow cells: Visual observation detected the close distance between two marmosets. Red cells: The system detected grooming by the intensity exceeding a threshold value.

References

1. Bala, P. C. *et al.* Automated markerless pose estimation in freely moving macaques with OpenMonkeyStudio. *Nat. Commun.* **11**, 4560 (2020) (DOI 10.1038/s41467-020-18441-5)

2. Yamazaki, Y. *et al.* Tool-use learning by common marmosets (Callithrix jacchus). *Exp. Brain Res.* **213**, 63–71 (2011) (DOI 10.1007/s00221-011-2778-9)

3. Miller, C. T. *et al.* Marmosets: A neuroscientific model of human social behavior. *Neuron* **90**, 219–233 (2016) (DOI 10.1016/j.neuron.2016.03.018)

4. Mitchell, J. F. & Leopold, D. A. The marmoset monkey as a model for visual neuroscience. *Neurosci. Res.* **93**, 20–46 (2015) (DOI 10.1016/j.neures.2015.01.008)

5. Roelfsema, P. R. & Treue, S. Basic neuroscience research with nonhuman primates: a small but indispensable component of biomedical research. *Neuron* **82**, 1200–1204 (2014) (DOI

574 10.1016/j.neuron.2014.06.003)

575 6. Burkart, J. M., Fehr, E., Efferson, C. & van Schaik, C. P. Other-regarding preferences in a

576 non-human primate: common marmosets provision food altruistically. *Proc. Natl Acad. Sci. U. S. A.*

577 **104**, 19762–19766 (2007) (DOI 10.1073/pnas.0710310104)

578 7. Abbott, D. H., Barnett, D. K., Colman, R. J., Yamamoto, M. E. & Schultz-Darken, N. J.

579 Aspects of common marmoset basic biology and life history important for biomedical research. *Comp.*

580 *Med.* **53**, 339–350 (2003)

581 8. Sasaki, E., *et al*. Generation of transgenic non-human primates with germline transmission.

582 *Nature* **459**, 523–527 (2009) (DOI 10.1038/nature08090)

583 9. Sato, K., *et al*. Generation of a nonhuman primate model of severe combined

584 immunodeficiency using highly efficient genome editing. *Cell Stem Cell* **19**, 127–138 (2016) (DOI

585 10.1016/j.stem.2016.06.003)

586 10. Sato, K., *et al*. A nonhuman primate model of familial Alzheimer's disease. *bioRxiv*,

587 2020.2008.2024.264259 (2020)

588 11. Kishi, N., Sato, K., Sasaki, E. & Okano, H. Common marmoset as a new model animal for

589 neuroscience research and genome editing technology. *Dev. Growth Differ.* **56**, 53–62 (2014) (DOI

590 10.1111/dgd.12109)

591 12. Okano, H., Miyawaki, A. & Kasai, Kasai K.. Brain/MINDS: brain-mapping project in Japan.

Philos. Trans. R. Soc. Lond. B Biol. Sci. **370** (2015) (DOI 10.1098/rstb.2014.0310)

13. network, B. I. C. C. A multimodal cell census and atlas of the mammalian primary motor cortex. *Nature* **598**, 86–102 (2021) (DOI 10.1038/s41586-021-03950-0)

14. Hong, W., *et al*. Automated measurement of mouse social behaviors using depth sensing, video tracking, and machine learning. *Proc. Natl Acad. Sci. U. S. A.* **112**, E5351–E5360 (2015) (DOI 10.1073/pnas.1515982112)

15. Redfern, W. S., *et al*. Automated recording of home cage activity and temperature of individual rats housed in social groups: the Rodent Big Brother project. *PLOS ONE* **12**, e0181068 (2017) (DOI 10.1371/journal.pone.0181068)

16. Sutoko, S., *et al.* Early identification of Alzheimer's disease in mouse models: application of deep neural network algorithm to cognitive behavioral parameters. *iScience* **24**, 102198 (2021) (DOI 10.1016/j.isci.2021.102198)

17. Hsu, A. I. & Yttri, E. A. B-SOiD, an open-source unsupervised algorithm for identification and fast prediction of behaviors. *Nat. Commun.* **12**, 5188 (2021) (DOI 10.1038/s41467-021-25420-x)

18. Karashchuk, P., Tuthill, J. C. & Brunton, B. W. The DANNCE of the rats: a new toolkit for 3D tracking of animal behavior. *Nat. Methods* **18**, 460–462 (2021) (DOI 10.1038/s41592-021-01110-w)

19. Zimmermann, C., Schneider, A., Alyahyay, M., Brox, T. & Diester, I. FreiPose: A deep

learning framework for precise animal motion capture in 3D spaces. *bioRxiv*, 2020.2002.2027.967620 (2020)

20. Yabumoto, T., *et al.* MarmoDetector: A novel 3D automated system for the quantitative assessment of marmoset behavior. *J. Neurosci. Methods* **322**, 23–33 (2019) (DOI 10.1016/j.jneumeth.2019.03.016)

21. Pereira, T. D., *et al.* SLEAP: A deep learning system for multi-animal pose tracking. *Nat. Methods* **19**, 486–495 (2022) (DOI 10.1038/s41592-022-01426-1)

22. Segalin, C., *et al*. The Mouse Action Recognition System (MARS) software pipeline for automated analysis of social behaviors in mice. *eLife* **10** (2021) (DOI 10.7554/eLife.63720)

23. Lauer, J., *et al*. Multi-animal pose estimation, identification and tracking with DeepLabCut. *Nat. Methods* **19**, 496–504 (2022) (DOI 10.1038/s41592-022-01443-0)

24. Mathis, A., *et al.* DeepLabCut: markerless pose estimation of user-defined body parts with deep learning. *Nat. Neurosci.* **21**, 1281–1289 (2018) (DOI 10.1038/s41593-018-0209-y)

25. de Chaumont, F., *et al*. Computerized video analysis of social interactions in mice. *Nat. Methods* **9**, 410–417 (2012) (DOI 10.1038/nmeth.1924)

26. Wiltschko, A. B., *et al*. Revealing the structure of pharmacobehavioral space through motion sequencing. *Nat. Neurosci.* **23**, 1433–1443 (2020) (DOI 10.1038/s41593-020-00706-3)

27. Sun, G., *et al.* DeepBhvTracking: A novel behavior tracking method for laboratory animals

based on deep learning. *Front. Behav. Neurosci.* **15**, 750894 (2021) (DOI 10.3389/fnbeh.2021.750894)

28. Schultz-Darken, N., Ace, L. & Chapter, A. H. 7. Behavior and behavioral management in *The Common Marmoset in Captivity and Biomedical Research* (eds Marini, R., Wachtman, L., Tardif, S., Mansfield, K. & Fox, J.) (Academic Press, 2019)

29. Redmon, J. & Yolov, F. A. An incremental improvement. *arXiv Preprint ArXiv:180402767*, (2018)

30. Matsumoto, J., *et al.* A 3D-video-based computerized analysis of social and sexual interactions in rats. *PLOS ONE* **8**, e78460 (2013) (DOI 10.1371/journal.pone.0078460)

31. Lyons, D. M., MacDonall, J. S. & Cunningham, K. M. A Kinect-based system for automatic recording of some pigeon behaviors. *Behav. Res. Methods* **47**, 1044–1054 (2015) (DOI 10.3758/s13428-014-0531-6)

32. Susanne, R. & Ann-Kathrin, O. Husbandry and management of New World species: marmosets and tamarins. *The Lab. Primate*, 145–162 (2005)

33. Kurotaki, Y. & Sasaki, E. Practical reproductive techniques for the common marmoset. *J. Mamm. Ova Res.* **34**, 3–12 (2017) (DOI 10.1274/032.034.0103)

34. Pines, M. K., Kaplan, G. & Rogers, L. J. A note on indoor and outdoor housing preferences of common marmosets (Callithrix jacchus). *Appl. Anim. Behav. Sci.* **108**, 348–353 (2007) (DOI 10.1016/j.applanim.2006.12.001)

35. Kobayashi, K. *et al.* Automated detection of mouse scratching behaviour using convolutional recurrent neural network. *Sci. Rep.* **11**, 658 (2021) (DOI 10.1038/s41598-020-79965-w) (Pubmed:33436724)

36. Bohnslav, J. P., *et al.* DeepEthogram, a machine learning pipeline for supervised behavior classification from raw pixels. *eLife* **10** (2021) (DOI 10.7554/eLife.63377)

37. Cohen, Y. *et al.* Automated annotation of birdsong with a neural network that segments spectrograms. *eLife* **11** (2022) (DOI 10.7554/eLife.63853)

38. Borges, C. R., Poyares, D., Piovezan, R., Nitrini, R. & Brucki, S. Alzheimer's disease and sleep disturbances: a review. *Arq. Neuro Psiquiatr.* **77**, 815–824 (2019) (DOI 10.1590/0004-282X20190149)

39. dos Santos, A. B., Kohlmeier, K. A. & Barreto, G. E. Are sleep disturbances preclinical markers of Parkinson's disease? *Neurochem. Res.* **40**, 421–427 (2015) (DOI 10.1007/s11064-014-1488-7)

40. Calapai, A., Cabrera-Moreno, J., Moser, T. & Jeschke, M. Flexible auditory training, psychophysics, and enrichment of common marmosets with an automated, touchscreen-based system. *Nat. Commun.* **13**, 1648 (2022) (DOI 10.1038/s41467-022-29185-9)

41. Zhou, J. *et al.* with Transposed Convolution for Rotating Machinery Fault Diagnosis Based on Deep Transfer Learning. *Shock Vib.* **2020**, 1–12 (2020)

42. Zhang, J., Xiao, W., Coifman, B. & Mills, J. P. Vehicle tracking and speed estimation from roadside lidar. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sensing* **13**, 5597–5608 (2020)

43. Rusu, R. B. Semantic 3D object maps for everyday manipulation in human living environments. *KI Künstliche Intell.* **24**, 345–348 (2010) (DOI 10.1007/s13218-010-0059-6)

Fig. 1

a

b

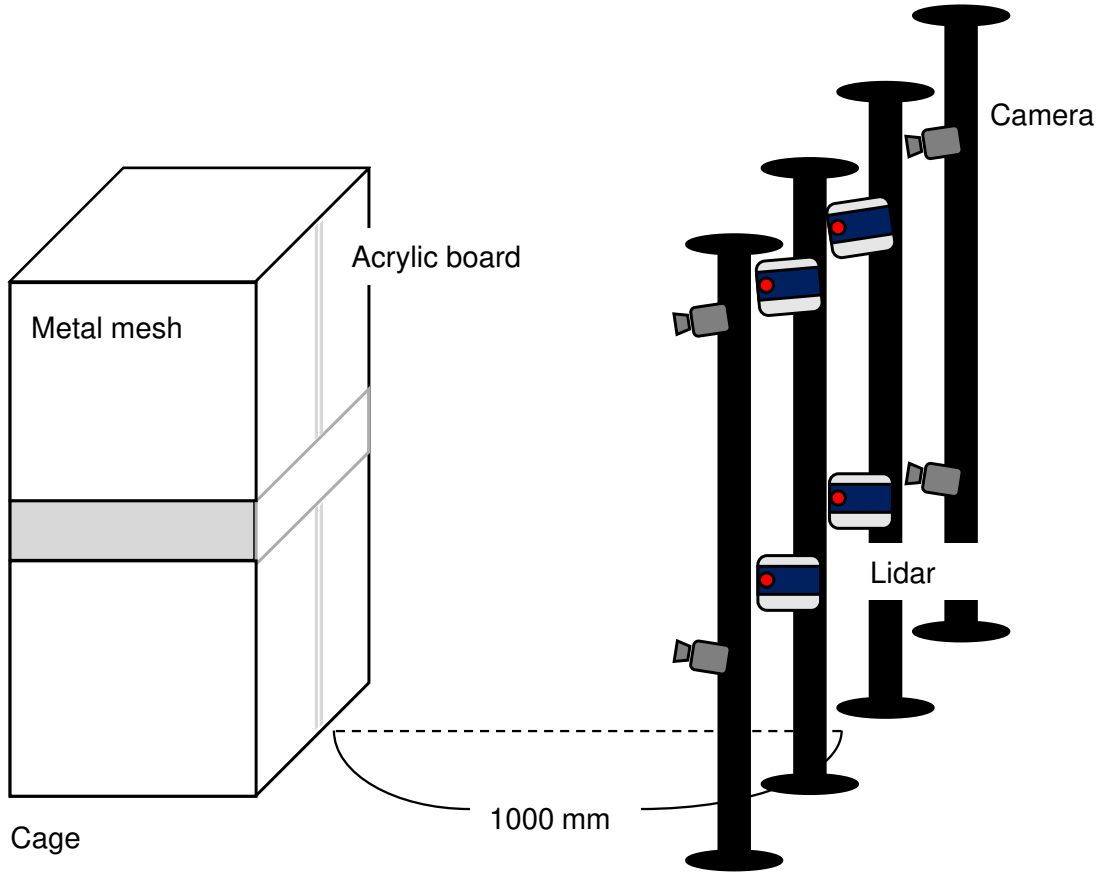


Fig. 2

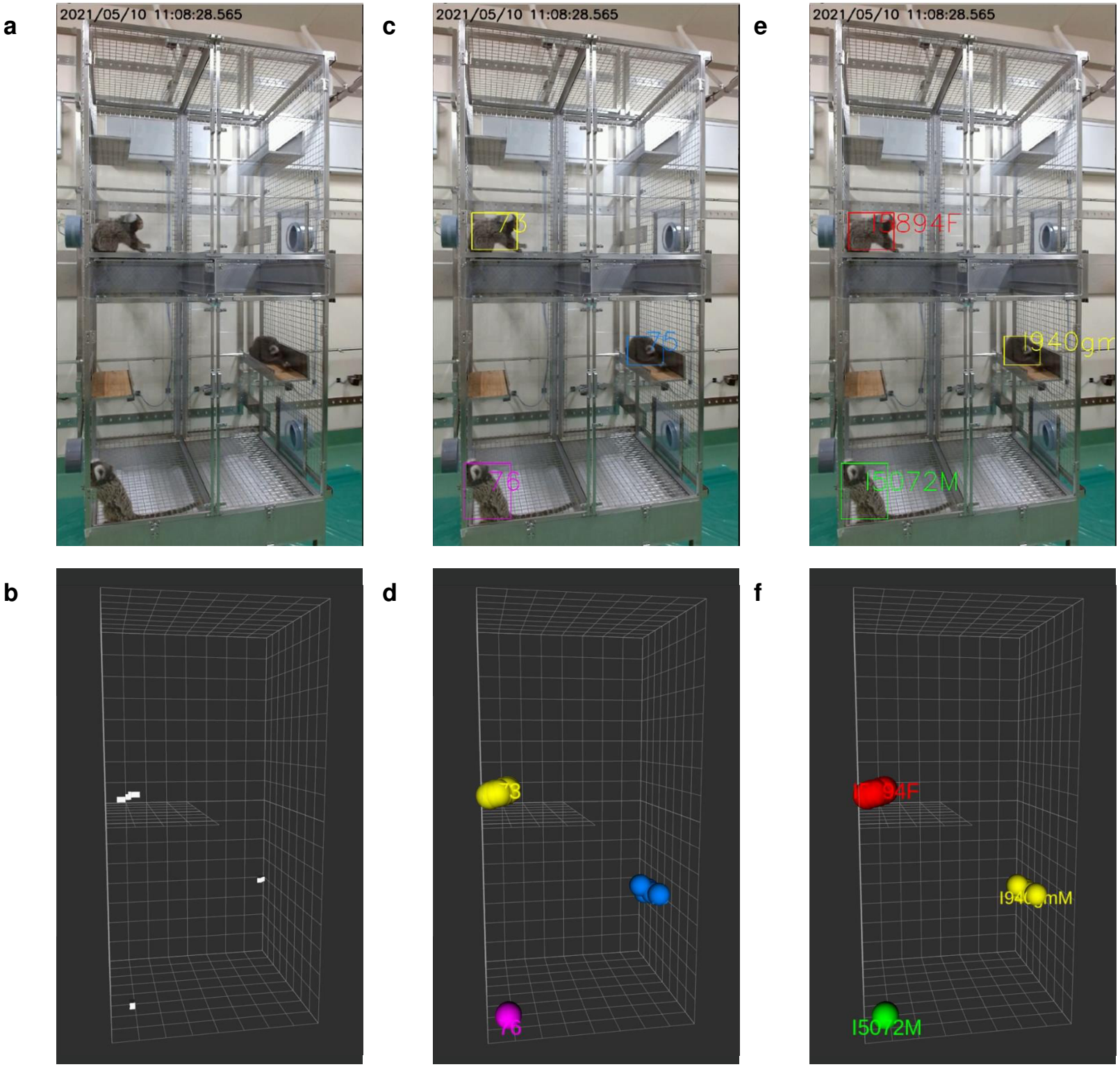

a
2021/05/10 11:08:28.565
c
2021/05/10 11:08:28.565
e
2021/05/10 11:08:28.565
b
d
f

Fig. 3

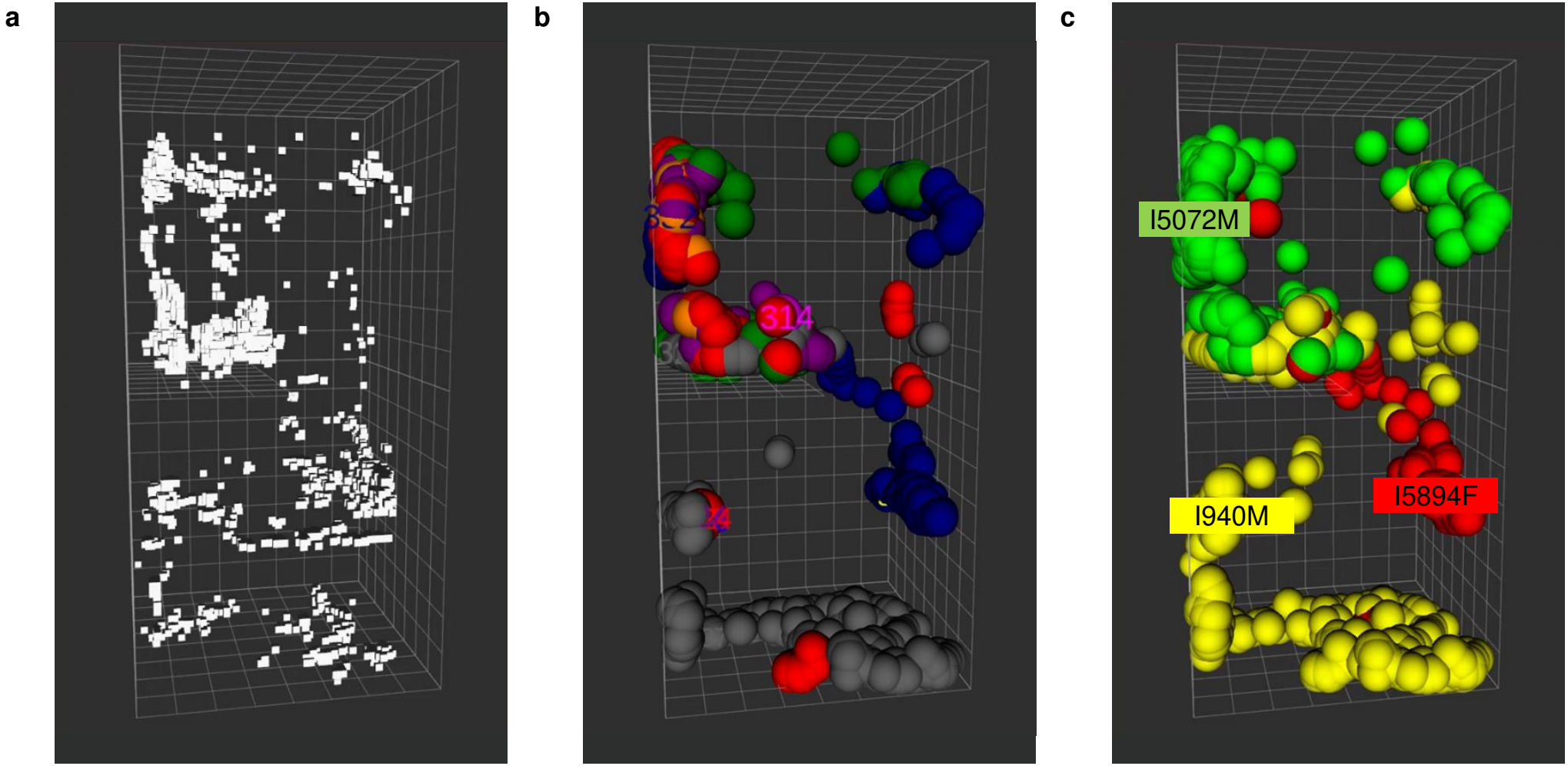


Fig. 4

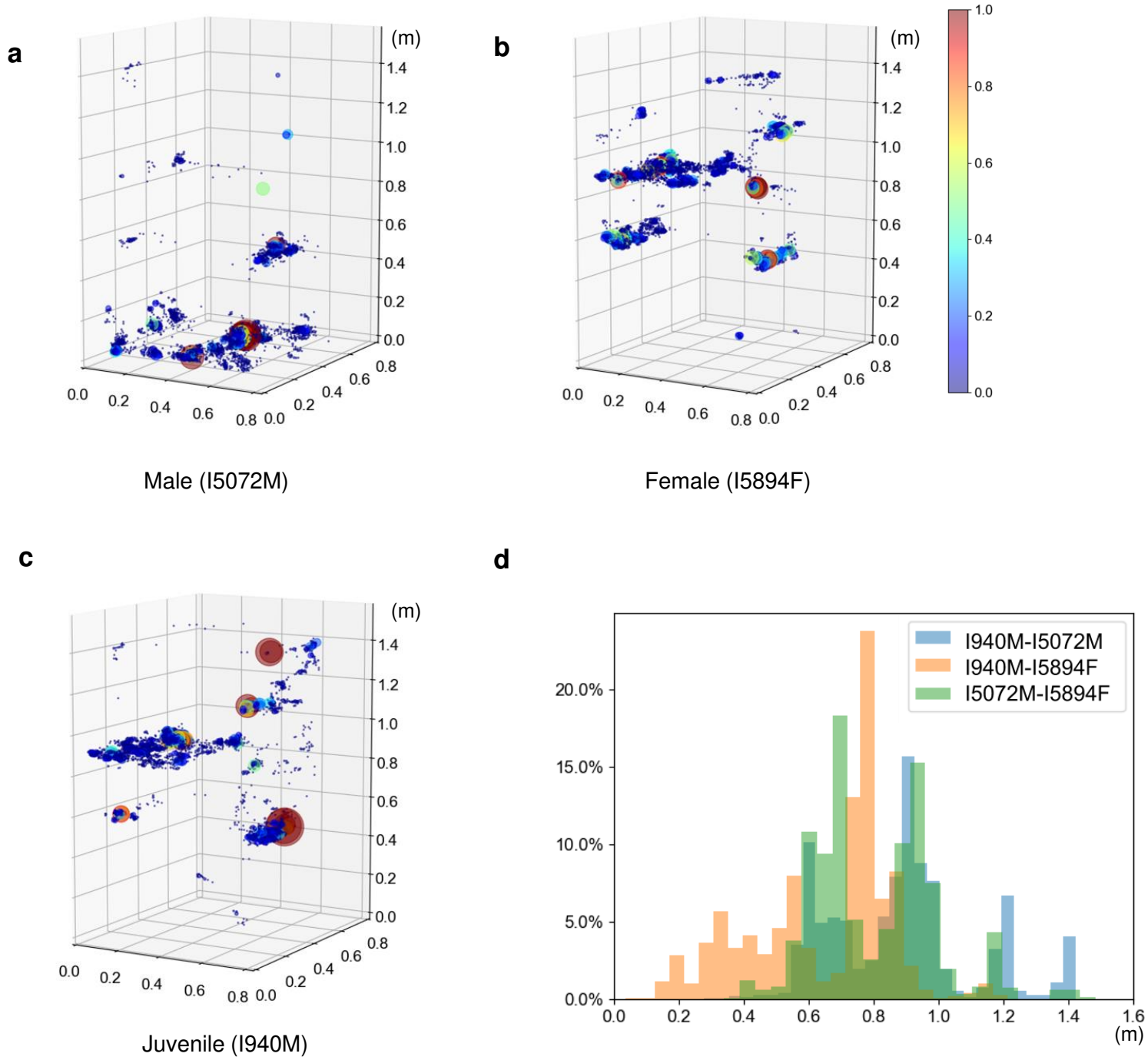
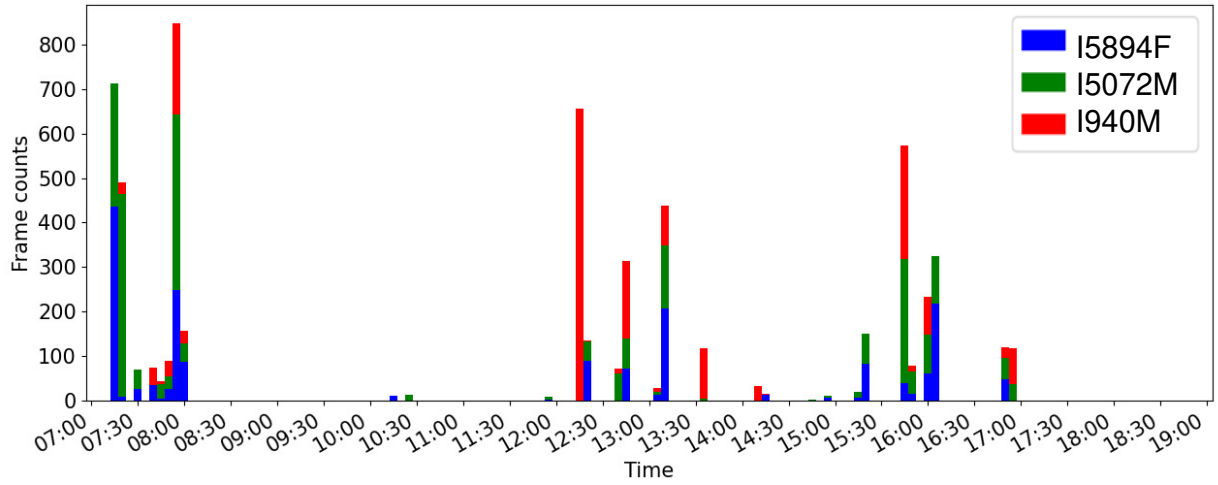

a
(m)
Male (I5072M)
b
(m)
Female (I5894F)
c
(m)
Juvenile (I940M)
d
I940M-I5072M
I940M-I5894F
I5072M-I5894F
(m)

Fig. 5

a


I5894F
I5072M
I940M
Frame counts
0
100
200
300
400
500
600
700
800
07:00
07:30
08:00
08:30
09:00
09:30
10:00
10:30
11:00
11:30
12:00
12:30
13:00
13:30
14:00
14:30
15:00
15:30
16:00
16:30
17:00
17:30
18:00
18:30
19:00
Time

b

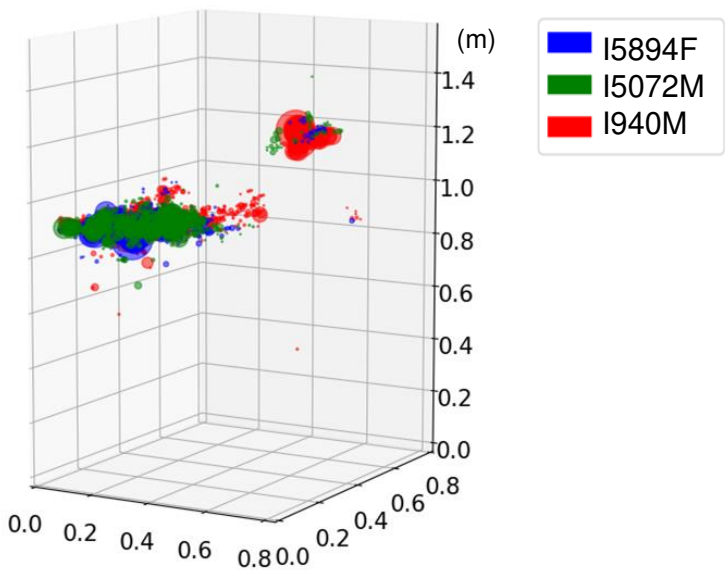
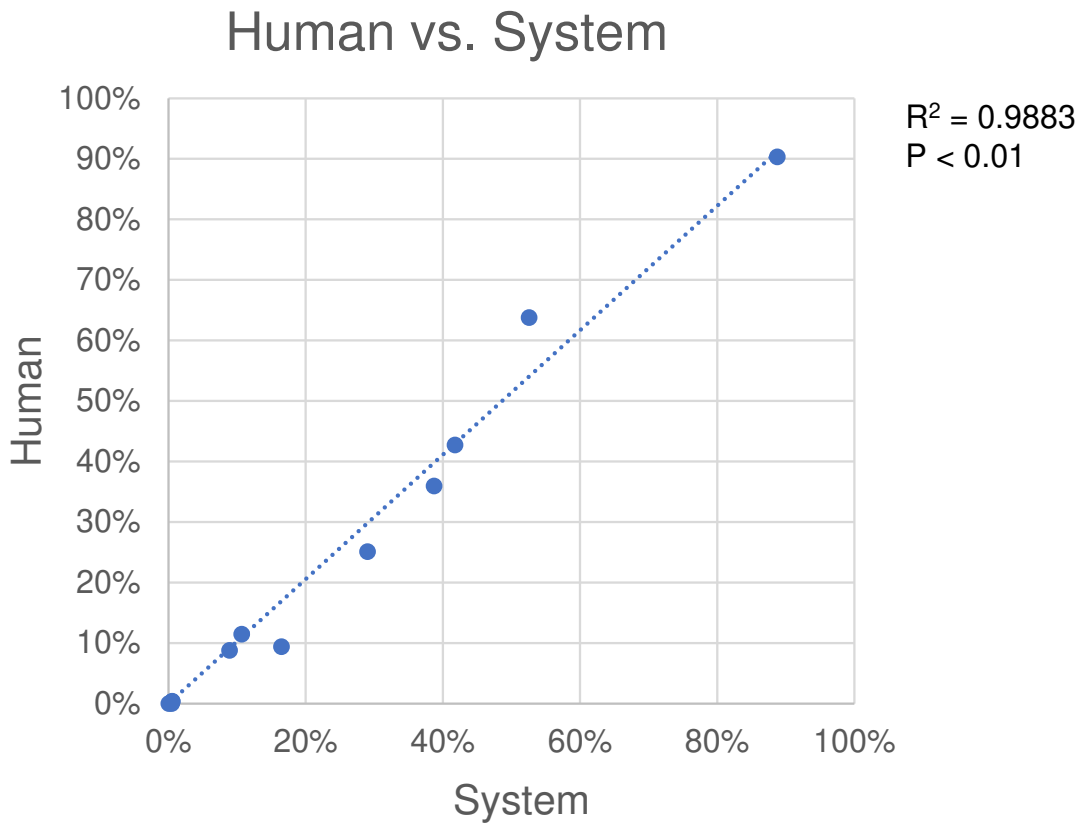

(m)
I5894F
I5072M
I940M
0.0
0.2
0.4
0.6
0.8
1.0
1.2
1.4

Fig. 6

a


Human vs. System
100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
Human
0%
20%
40%
60%
80%
100%
System
R^2 = 0.9883
P < 0.01

b

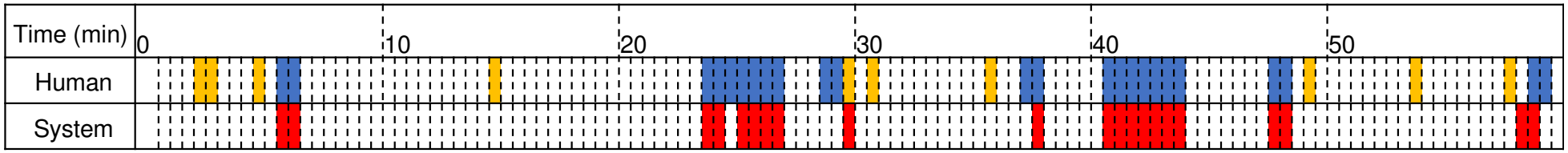

Time (min)
0
10
20
30
40
50
Human
System

Table 1. Classification accuracy of face identification in five marmosets aged 1 to 2 years.

	Animal ID				
	I774	I6695	I6708	I6677	I893
Specificity	100.0%	99.8%	99.8%	99.8%	99.8%
Accuracy	99.8%	99.6%	99.4%	99.8%	99.8%
Recall	99.0%	99.0%	98.0%	100.0%	100.0%
Precision	100.0%	99.0%	99.0%	99.0%	99.0%
F-Measure	99.5%	99.0%	98.5%	99.5%	99.5%

Table 2. Classification accuracy of face identification in a marmoset family.

	Male (I5072M)	Female (I5894F)	Juvenile (I940M)	Unknown
Specificity	98.7%	99.5%	99.7%	99.5%
Accuracy	98.7%	98.6%	99.2%	99.5%
Recall	98.8%	96.1%	97.6%	99.6%
Precision	96.1%	98.4%	99.1%	98.5%
F-Measure	97.4%	97.3%	98.3%	99.0%

Table 3. Comparison of visual observations and the system analysis of time spent in specific locations in the cage for an hour

	Male (I5072M)		Female (I5894F)		Juvenile (I940M)	
	Human	System	Human	System	Human	System
Upper bed	0.4%	0.6%	10.7%	16.5%	9.4%	16.5%
Middle floor	0.3%	0.6%	38.8%	52.6%	63.8%	52.6%
Lower bed	8.8%	8.9%	41.8%	29.1%	25.1%	29.1%
Lower floor	90.3%	88.8%	0.5%	0.5%	0.0%	0.1%
Other	0.2%	1.11%	8.2%	8.2%	1.7%	1.7%