Journal of Visualized Experiments

A Step-By-Step Implementation of DeepBehavior, Deep Learning Toolbox for the Automated Behavior Analysis

--Manuscript Draft--

Article Type:	Invited Methods Article - JoVE Produced Video
Manuscript Number:	JoVE60763R1
Full Title:	A Step-By-Step Implementation of DeepBehavior, Deep Learning Toolbox for the Automated Behavior Analysis
Section/Category:	JoVE Behavior
Keywords:	deep learning; behavior analysis; Convolutional Neural Nets; machine learning; Kinematic Analysis; Automated Analysis; animal behavior; Human Behavior; Reaching tasks; Image Data; Video Data; 3D kinematics
Corresponding Author:	Ahmet Arac, MD UCLA Los Angeles, CA UNITED STATES
Corresponding Author's Institution:	UCLA
Corresponding Author E-Mail:	AArac@mednet.ucla.edu
Order of Authors:	Sanjay Shukla
	Ahmet Arac, MD
Additional Information:	
Question	**Response**
Please indicate whether this article will be Standard Access or Open Access.	Standard Access (US$2,400)
Please indicate the **city, state/province, and country** where this article will be **filmed**. Please do not use abbreviations.	Los Angeles, CA, USA

TITLE:

A Step-By-Step Implementation of DeepBehavior, Deep Learning Toolbox for the Automated Behavior Analysis

AUTHORS AND AFFILIATIONS:
Sanjay Shukla[1], Ahmet Arac[1]

[1]Department of Neurology, David Geffen School of Medicine, University of California, Los Angeles, Los Angeles, CA, United States

Corresponding Author:
Ahmet Arac (aarac@mednet.ucla.edu)

Email Addresses of Co-authors:
Sanjay Shukla (sanjayshukla@ucla.edu)

KEYWORDS:
deep learning, behavior analysis, convolutional neural nets, machine learning, kinematic analysis, automated analysis, animal behavior, human behavior, reaching tasks, image data, video data, 3D kinematics

SUMMARY:
The purpose of this protocol is to utilize pre-built convolutional neural nets to automate behavior tracking and perform detailed behavior analysis. Behavior tracking can be applied to any video data or sequences of images and is generalizable to track any user-defined object.

ABSTRACT:
Understanding actions is the first step to truly comprehend neural mechanisms in the brain that drive behavior. Traditional behavioral analysis methods often do not capture the richness inherent to natural behavior. Here, we provide detailed step-by-step instructions with visualizations of our recent methodology, DeepBehavior. The DeepBehavior toolbox uses deep learning frameworks integrated with convolutional neural networks to rapidly process and analyze behavioral videos. This protocol demonstrates three different frameworks for single object detection, multiple object detection, and three-dimensional (3D) human joint pose tracking. These frameworks return Cartesian coordinates of the object of interest for each frame of the behavior video. Data collected from the DeepBehavior toolbox contain much more detail than traditional behavior analysis methods and provide detailed insights to the behavior dynamics. DeepBehavior quantifies behavioral tasks in a robust, automated, and precise way. Following the identification of behavior, postprocessing code is provided to extract information and visualizations from the behavioral videos.

INTRODUCTION:

A detailed analysis of behavior is the key to understanding the brain and behavior relationships. There have been many exciting advances in methodologies for recording and manipulating neuronal populations with high temporal resolution. However, behavior analysis methods have not developed at the same rate and are limited to indirect measurements and a reductionist approach[1]. Recently, deep learning-based methods have been developed to perform automated and detailed behavior analysis[2-5]. This protocol provides a step-by-step implementation guide for the DeepBehavior toolbox.

Traditional behavioral analysis methods often include manually labeling data from multiple evaluators, leading to variance in how experimenters define a behavior[6]. Manual labeling of the data requires time and resources that increase disproportionately to the amount of data collected. Moreover, manually labelling data reduces the behavior outcomes into categorical measurements that do not capture the richness of the behavior and are more subjective. Thus, the current traditional methods may be limited in capturing the details in the natural behaviors.

The DeepBehavior toolbox presents a precise, detailed, highly temporal, and automated solution to using deep learning for behavioral analysis. Deep learning has quickly become accessible to all with open-source tools and packages. Convolutional neural networks (CNNs) are proven to be highly effective in object recognition and tracking tasks[7,8]. Using modern day CNNs and high-performance graphics-processing-units (GPUs), large images and videos can be processed quickly with high precision[7,9-11]. In DeepBehavior, there are three different convolutional neural net architectures, TensorBox, YOLOv3, and OpenPose[2].

The first framework, TensorBox, is a versatile framework that incorporates many different CNN architectures for object detection[12]. TensorBox is best suited for detecting only one object class per image. The resulting outputs are bounding boxes of the object of interest (**Figure 1**) and the Cartesian coordinates of the bounding box.

The second CNN framework is YOLOv3, which stands for "You Only Look Once"[13]. YOLOv3 is advantageous when there are multiple objects of interest that must be tracked separately. The output of this network includes the bounding box with the associated object label class as well as the bounding box Cartesian coordinates of the object in the video frame (**Figure 2**).

The previous two frameworks are valuable for generalized behavioral data collected from standard laboratory experiments in animal subjects. The last CNN framework is OpenPose[14-16], which is used for human joint pose estimation. OpenPose detects human body, hand, facial, and foot key points on images. The outputs of the framework are labeled images of the human subject as well as the coordinates of all 25 key points in the body and 21 key points of each hand (**Figure 3**).

This detailed step-by-step guide for the implementation of our recently developed open-source DeepBehavior toolbox employs state of the art convolutional neural nets to track animal behavior (e.g., movement of a paw) or human behavior (e.g., reaching tasks). By tracking behavior, useful kinematics can be derived from the actions, such as position, velocity, and acceleration. The protocol explains the installation of each CNN architecture, demonstrates how to create training datasets, how to train the networks, how to process new videos on the trained network, how to extract the data from the network on the new videos, and how to postprocess the output data to make it useful for further analysis.

PROTOCOL:

1. GPU and Python set-up

1.1. GPU software

1.1.1. Set up the computer for deep learning applications. Ensure that GPU-appropriate software and drivers are installed, which can be found on the GPU's respective website (see **Table of Materials**).

1.2. Python 2.7 installation

1.2.1**.** Use the following command line prompt on the machine to install Python.
sudo apt-get install Python-pip Python-dev Python-virtualenv

2. TensorBox

2.1. Setting up TensorBox

2.1.1. Create the Virtual Environment for TensorBox using the commands below.
cd ~
virtualenv --system-site-packages ~/tensorflow

NOTE: '*~/tensorflow*' is the name of the environment and is arbitrary.

2.1.2. Activate the environment with the following command.
source ~/tensorflow/bin/activate

2.2. Install TensorBox.

2.2.1. Use GitHub to clone TensorBox from http://github.com/aarac/TensorBox and install it on the machine as well as on additional dependencies.
cd ~
git clone http://github.com/aarac/TensorBox
cd tensorbox
pip install -r requirements.txt

2.3. Label data.

2.3.1. Create a folder of behavior images. Use open-source tools such as ffmpeg to convert videos to individual frames. Label at least 600 images from a wide distribution of behavior frames for training. Put these images in a folder.

2.3.2. Launch the labeling graphical user interface.
Python make_json.py <path to image folder> train.json

2.3.3. To label an image, click the top left corner of the object of interest (i.e., paw) first and then click the bottom right corner of the object of interest (**Figure 4**). Check that the bounding box captures the entire object of interest. Press '**Undo**' to relabel the same image or press '**Next**' to move on to the next frame.

2.4. Train TensorBox.

2.4.1. Link training images to a network hyperparameters file. To do so, within the TensorBox folder, open the following folder in a text editor: */tensorbox/hypes/overfeat_rezoom.json*. Navigate to the attribute under *data* named *train_idl* and replace the file path from *./data/brainwash/train_boxes.json* to *train.json*. Save the changes to the file.

2.4.2. Begin the training script.
cd ~/tensorbox
Python train.py --hypes hypes/overfeat_rezoom.json --gpu 0 --logdir output

NOTE: The network will then begin training for 600,000 iterations. In the output folder, the resulting trained weights of the convolutional neural network will be generated.

2.5. Perform prediction on new images.

2.5.1. Perform image labeling.
cd ~/tensorbox
Python label_images.py --folder <path to image folder> --weights output/overfeat_rezoom_<timestamp>/save.ckpt-600000 --hypes /hypes/overfeat_rezoom.json --gpu 0

2.5.2. Then get the coordinates of bounding boxes.
cd ~/tensorbox
Python predict_images_to_json.py --folder <path to image folder> --weights output/overfeat_rezoom_<timestamp>/save.ckpt-600000 --hypes /hypes/overfeat_rezoom.json --gpu 0

2.6 MATLAB postprocessing for TensorBox.

NOTE: Additional MATLAB code has been provided to extract kinematics and visualizations of the coordinates using the resulting JSON coordinate file from the model.

2.6.1. Run the "Process_files_3Dreaching_mouse.m" script for 3D kinematic analysis of a single food pellet reaching task.

3. YOLOv3

3.1. Install YOLOv3.

```
cd ~
git clone https://github.com/aarac/darknet
cd darknet
make
```

3.2 Perform labeling of training data using Yolo_mark.

```
cd ~
git clone https://github.com/aarac/Yolo_mark
cd ~/Yolo_Mark
cmake
make
```

3.2.1. Place the training images in the *~/Yolo_mark/data/obj* folder.

```
chmod +x ./Linux_mark.sh
 ./Linux_mark.sh
```

3.2.2. Label the images one by one in the graphical user interface (**Figure 5**). The recommended number of images is approximately 200.

3.3. Training YOLOv3

3.3.1. Set up the configuration file.

```
cd ~/Yolo_mark
scp -r ./data ~/darknet
cd ~/darknet/cfg
cp yolov3.cfg yolo-obj.cfg
```

3.3.2. Modify the configuration file. To do so, open the yolo-obj.cfg folder and modify the following lines: batch=64, subdivision=8, classes=(# of class to detect), and for each convolutional layer before a yolo layer change the filter=(classes+5)x3. Details on these changes can be found at https://github.com/aarac/darknet/blob/master/README.md

3.3.3. Download network weights from https://www.dropbox.com/s/613n2hwm5ztbtuf/darknet53.conv.74?dl=0. Place the downloaded weight file into *~/darknet/build/x64*

3.3.4. Run the training algorithm.

cd ~/darknet
./darknet detector train data/obj.data cfg/yolo-obj.cfg darknet53.conv.74

3.3.5. YOLOv3 evaluation: After the training is complete with a set number of iterations (*ITERATIONNUMBER)*, view them by.
./darknet detector test data/obj.data cfg/yolo-obj.cfg backup/yolo-obj_ITERATIONNUMBER.weights <IMAGE>.jpg

3.4. Make predictions regarding new videos and get coordinates.

3.4.1. Use this command to obtain the coordinates of the labels in the new video.
./darknet detector demo data/obj.data cfg/yolo-obj.cfg backup/yolo-obj_ITERATIONNUMBER.weights VIDEO.avi -ext_output <VIDEO.avi> FILENAME.txt

3.5 YOLOv3 post processing in MATLAB

3.5.1. Take the FILENAME.txt file to MATLAB and run the "Process_socialtest_mini.m" script for two mice social interaction test. See results in **Figure 2.**

4. OpenPose

NOTE: OpenPose is ideal to track multiple body parts in a human subject. The set up and installation processes are very similar to the previous two frameworks. However, there is no training step, because the network is already trained on human data.

4.1. OpenPose Installation

4.1.1. Navigate to https://github.com/aarac/openpose and follow the installation instructions.

4.2. Process the video.
./build/examples/openpose/openpose.bin --num_gpu 0 --video VIDEONAME.avi --net_resolution "1312x736" --scale_number 4 --scale_gap 0.25 --hand --hand_scale_number 6 --hand_scale_range 0.4 --write_json JSONFOLDERNAME --write_video RESULTINGVIDEONAME.avi

NOTE: Here the –net_resolution, --scale_number, --scale_gap, --hand_scale_number and –hand_scale_range handles can be omitted if a high precision detection is not needed (this decreases the processing time).

4.3. Perform OpenPose postprocessing.

4.3.1. In the MATLAB folder, use *'process_files_human3D.m'* script to run the code after adding the appropriate folder containing json files from cameras 1 and 2, as well as the calibration file. This will create a "cell" file with all the 3D poses of the joints and will also make a movie of the

3D skeletal view. For camera calibration, follow the instructions at this link: http://www.vision.caltech.edu/bouguetj/calib_doc/.

REPRESENTATIVE RESULTS:

When the protocol is followed, the data for each network architecture should be as follows: TensorBox outputs a bounding box around the object of interest. In our example, we used videos from a food pellet reaching task, and labeled the right paws to track their movement. As seen in **Figure 1**, the right paw was detected in different positions in both the front view and side view cameras. After postprocessing with camera calibration, 3D trajectories of the reach were obtained (**Figure 1B**).

In Yolov3, because there are multiple objects, the output is also multiple bounding boxes. As seen in **Figure 2B,** there were multiple bounding boxes around the objects of interest. These can be parts of the body.

In OpenPose, the network detected the joint positions as seen in **Figures 3A** and **3B**. After postprocessing with camera calibration, a 3D model of the subject was created (**Figure 3C**).

In conclusion, these representative results showcase the rich details of behavior that can be captured using the DeepBehavior toolbox.

FIGURE AND TABLE LEGENDS:

Figure 1: Bounding boxes with TensorBox seen on the paws of video frames during a reaching task in mice. Adapted with permission from Arac et al.[2].

Figure 2: Bounding boxes with Yolov3 seen on the regions of interest in video frames during a two mice social interaction test. (**A**) Raw image. (**B**) Analyzed image. Adapted with permission from Arac et al.[2].

Figure 3: Human pose detection with OpenPose. Pose shown in two camera views **(A**, **B)**. A 3D model created from these two images **(C)**. Adapted with permission from Arac et al.[2].

Figure 4: TensorBox's make_json GUI used to label training data.

Figure 5: GUI of Yolo_Mark to label images in a format acceptable for Yolov3.

DISCUSSION:

Here, we provide a step-by-step guide for implementation of DeepBehavior, our recently developed deep learning-based toolbox for animal and human behavior imaging data analysis[2]. We provide detailed explanations for each step of the installation of the frameworks for each network architecture and provide links for installation of the open-source requirements to be able to run these frameworks. We demonstrate how to install them, how to create training data, how to train the network, and how to process new video files on the trained network. We also provide the postprocessing code to extract the basic necessary information needed for further analysis.

For single object detection, we recommend using TensorBox. If the goal is to track multiple objects at once, we recommend using YOLOv3. Finally, to obtain human kinematic data, we recommend using OpenPose. In this protocol we have shown that deep learning methods are able to process hundreds of thousands of frames while tracking objects with a high degree of precision. Using the postprocessing code provided, we can derive meaningful ways of analyzing the tracked behavior of interest. This provides a more detailed way of capturing behavior. It also provides an automated, robust way of defining behavior that is generalizable to many different types of behavioral tasks.

It is quite common to get a 'ModuleNotFoundError' when starting with a new virtual environment or code that has been downloaded from the internet. If this occurs, open up the terminal, activate the source environment, and type *'pip install <missing module name>'.* If the problem persists, you will need to check your Python version as well as other dependency packages.

Limitations to this technique include the technical troubleshooting necessary to properly set up GPU processing units compatible with open-source code. It is advantageous to have past programming experience within a Linux environment to properly set up the necessary project dependencies and environments that are compatible with the computer's hardware.

We demonstrate the DeepBehavior toolbox installations and processing in a Linux environment. This toolbox can also be run on a Windows machine with GPUs, however. For Macintosh users, Yolov3 and TensorBox are fully compatible for GPU usage while OpenPose is only compatible with CPU usage.

Using deep learning methods for imaging data analysis is a very efficient way to automate behavior analysis. In comparison to traditional behavior analysis methods, DeepBehavior captures much more information to quantify, automate, and evaluate behavior at a more precise and temporally detailed way. With further advances in the deep learning field, the utilization and extent of use of this technology in behavior analysis will likely continue to improve. The applications of DeepBehavior can be expanded beyond the demonstrated reaching tasks to identify objects of interest in any behavioral images. In this protocol, we provide detailed instructions to implement three neural networks for behavior analysis. With this kind of automated and unbiased behavior analysis methods the neuroscience field will hopefully be able to do more detailed behavior analysis.

ACKNOWLEDGMENTS:
We would like to thank Pingping Zhao and Peyman Golshani for providing the raw data for two-mouse social interaction tests used in the original paper[2]. This study was supported by NIH NS109315 and NVIDIA GPU grants (AA).

DISCLOSURES:
The authors have nothing to disclose.

REFERENCES:

1. Krakauer, J. W., Ghazanfar, A. A., Gomez-Marin, A., MacIver, M. A., Poeppel, D. Neuroscience Needs Behavior: Correcting a Reductionist Bias. *Neuron.* **93** (3), 480–490 (2017).
2. Arac, A., Zhao, P., Dobkin, B. H., Carmichael, S. T., Golshani, P. DeepBehavior: A Deep Learning Toolbox for Automated Analysis of Animal and Human Behavior Imaging Data. *Frontiers in Systems Neuroscience.* **13**, 20 (2019).
3. Pereira, T. D. et al. Fast animal pose estimation using deep neural networks. *Nature Methods.* **16** (1), 117–125 (2019).
4. Mathis, A. et al. DeepLabCut: markerless pose estimation of user-defined body parts with deep learning. *Nature Neuroscience.* **21** (9), 1281–1289 (2018).
5. Stern, U., He, R., Yang, C. H. Analyzing animal behavior via classifying each video frame using convolutional neural networks. *Science Reports*. **5**, 14351 (2015).
6. Tinbergen, N. On aims and methods of ethology. *Zeitschrift für Tierpsychologie.* 20, 410:33 (1963).
7. LeCun, Y., Bengio, Y., Hinton, G. Deep learning. *Nature.* **521** (7553), 436–44 (2015).
8. Zhao, Z., Zheng, P., Xu, S., Wu, X. Object Detection With Deep Learning: A Review. *IEEE Transactions on Neural Networks and Learning Systems.* 1–21 (2019).
9. He, K., Zhang, X., Ren, S., Sun, J. Deep Residual Learning for Image Recognition. eprint *arXiv:*151203385. 2015:arXiv:1512.03385.
10. Krizhevsky, A., Sutskever, I, Hinton, G. E. ImageNet classification with deep convolutional neural networks. Proceedings of the 25th International Conference on Neural Information Processing Systems - Volume 1; Lake Tahoe, Nevada. 2999257: Curran Associates Inc.; 2012. p. 1097–1105.
11. Szegedy, C. et al., eds. Going deeper with convolutions. 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR); 2015 7–12 June 2015.
12. Stewart, R., Andriluka, M., Ng, A. Y., eds. End-to-End People Detection in Crowded Scenes. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR); 2016 27–30 June 2016.
13. Redmon, J., Farhadi, A. YOLOv3: An Incremental Improvement. eprint arXiv:180402767. 2018:arXiv:1804.02767.
14. Cao, Z., Simon, T., Shih-En, W., Sheikh, Y. Realtime Multi-Person 2D Pose Estimation using Part Affinity Fields. *arXiv.* 2017;1611(08050v2).
15. Simon, T., Joo, H., Matthews, I., Sheikh, Y. Hand Keypoint Detection in Single Images using Multiview Bootstrapping. eprint *arXiv:*170407809. 2017:arXiv:1704.07809.
16. Wei, S-E., Ramakrishna, V., Kanade, T., Sheikh, Y. Convolutional Pose Machines. eprint *arXiv:*160200134. 2016:arXiv:1602.00134.

Figure 1

Click here to access/download;Figure;figure1 copy.pdf

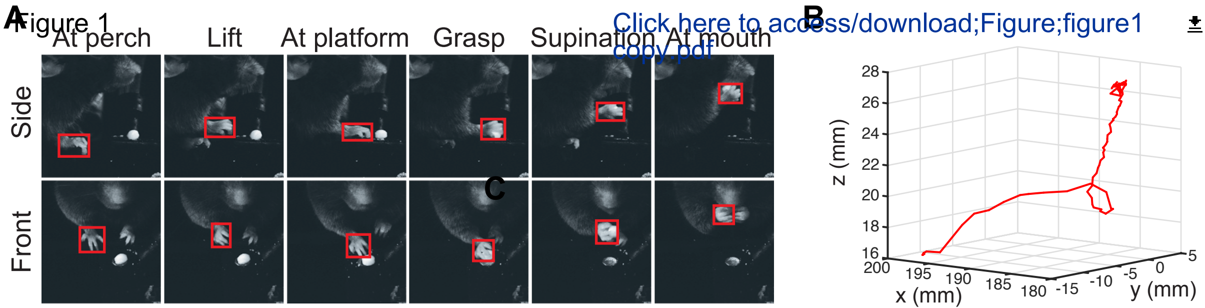


Figure 2

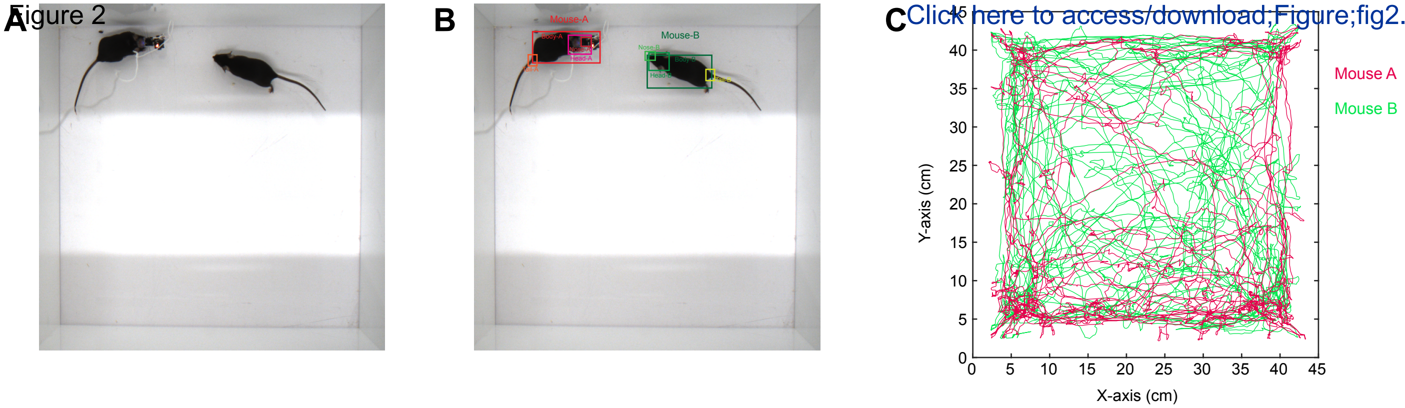
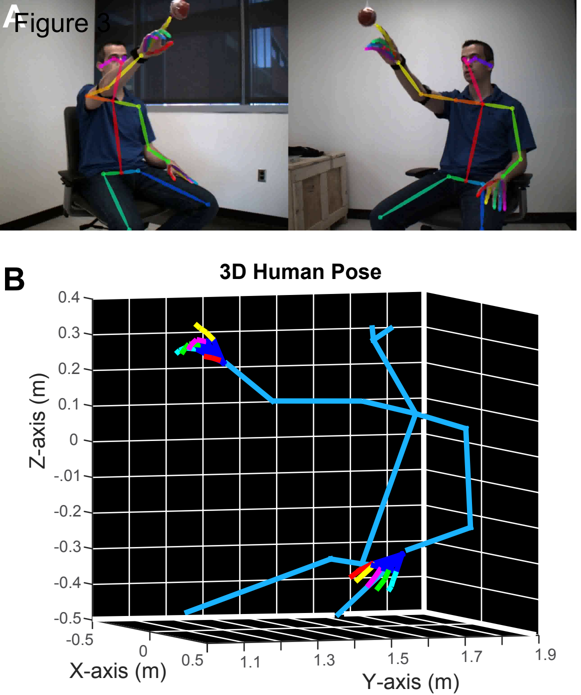

A
B
Mouse-A
Body-A
Nose-A
Head-A
Tail-A
Mouse-B
Nose-B
Body-B
Head-B
Tail-B
C
Mouse A
Mouse B
Y-axis (cm)
X-axis (cm)
0
5
10
15
20
25
30
35
40
45

Figure 3



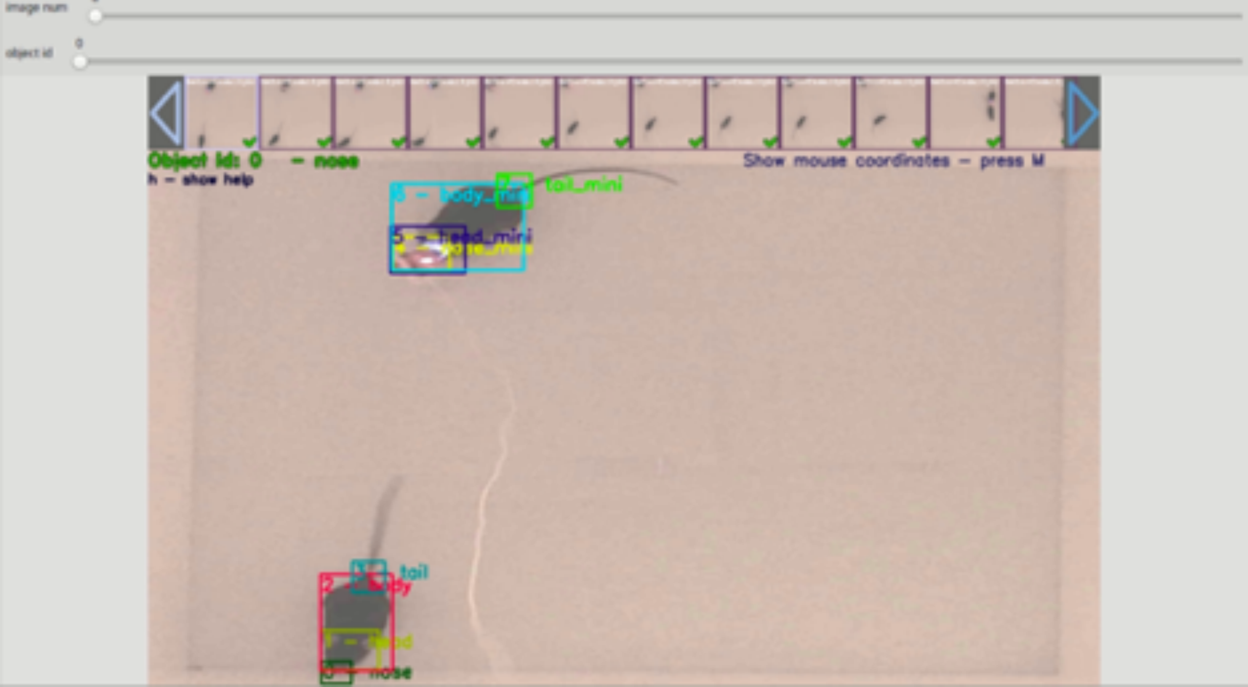
A
B
3D Human Pose
Z-axis (m)
0.4
0.3
0.2
0.1
0
-.01
-0.2
-0.3
-0.4
-0.5
X-axis (m)
-0.5
0
0.5
Y-axis (m)
1.1
1.3
1.5
1.7
1.9

Figure 4

Click here to access/download;Figure;Figure4.png

Figure 5

Click here to access/download;Figure;fig5.pdf



Name of Material/ Equipment	Company	Catalog Number	Comments/Description
CUDA v8.0.61	NVIDIA	n/a	GPU Software
MATLAB R2016b	Mathworks	n/a	Matlab
Python 2.7	Python	n/a	Python Version
Quadro P6000	NVIDIA	n/a	GPU Processor
Ubuntu v16.04	Ubuntu	n/a	Operating System

UNIVERSITY OF CALIFORNIA, LOS ANGELES **UCLA**

BERKELEY • DAVIS • IRVINE • LOS ANGELES • RIVERSIDE • SAN DIEGO • SAN FRANCISCO SANTA BARBARA • SANTA CRUZ

DEPARTMENT OF NEUROLOGY
710 WESTWOOD PLAZA
REED NEUROLOGICAL INSTITUTE
LOS ANGELES, CA 90095
TELEPHONE: (310) 794-1195
FAX: (310) 206-0987

November 11, 2019

JoVE
1 Alewife Center
Suite 200
Cambridge MA 02140

Dear Editor,

Please find attached a revision of our manuscript titled “Step-by-step implementation of DeepBehavior: A deep learning toolbox for automated behavior analysis” for consideration for publication in an appropriate JoVE journal, preferably JoVE Behavior or Neuroscience.

We have made all the recommended changes and please see our point-by-point responses to Editorial and Reviewer comments below.

Thank you for your consideration.

Sincerely,

Ahmet Arac, M.D.
Assistant Professor
Department of Neurology
UCLA
Los Angeles, CA

RESPONSES:

Editorial comments:

Changes to be made by the Author(s):

1. Please take this opportunity to thoroughly proofread the manuscript to ensure that there are no spelling or grammar issues. The JoVE editor will not copy-edit your manuscript and any errors in the submitted revision may be present in the published version.

Response: We have revised the manuscript accordingly.

2. Please revise the title to be more concise and avoid punctuations.

Response: We have now shortened the title and made it more concise.

3. Please remove all commercial language from your manuscript and use generic terms instead. All commercial products should be sufficiently referenced in the Table of Materials and Reagents.
For example: NVIDIA, CUDA versions, CUDNN, etc.
Response: We have revised it accordingly.

4. All steps in the protocol should be a numbered action step. Please adjust the numbering of the Protocol to follow the JoVE Instructions for Authors. For example, 1 should be followed by 1.1 and then 1.1.1 and 1.1.2 if necessary. Please refrain from using bullets or dashes.
Response: We have numbered the sections accordingly.

5. Please ensure that all text in the protocol section is written in the imperative tense as if telling someone how to do the technique in a step wise manner from beginning to the end (e.g., “Do this,” “Ensure that,” etc.). The actions should be described in the imperative tense in complete sentences wherever possible. Avoid usage of phrases such as “could be,” “should be,” and “would be” throughout the Protocol. Any text that cannot be written in the imperative tense may be added as a “Note.”
Response: We have ensured to use imperative tense.

6. The Protocol should be made up almost entirely of discrete steps without large paragraphs of text between sections. Please simplify the Protocol so that individual steps contain only 2-3 actions per step.
Response: We have made the steps more discrete.

7. Please revise the protocol text to avoid the use of any personal pronouns in the protocol (e.g., "we", "you", "our" etc.).
Response: We have removed the personal pronouns.

8. Please add more details to your protocol steps. Please ensure you answer the “how” question, i.e., how is the step performed?
Response: We have added additional details to make it clearer.

9. All software steps must be explained explicitly. Please include GUIs, button clicks in the softwares, command lines etc.
Response: We have included and revised Figure-4 for this.

10. Please make the protocol subheadings as installations, pretraining preparations, data collection (how are the human and animal data collected, which behavior/s is/are recorded, how many videos are made), validating the framework, configuring the experiment, executing the experiment, data collection, data analysis, etc. So, the protocol should show how you perform your experiment in a step wise manner providing all specific actions in detail. Please do not generalize. Maybe take the examples described in the result section and build up the whole protocol on it.

Response: As this manuscript is focused on the analysis of behavior imaging data (rather than the acquisition), we have made each step very clear.

11. 2: is tensorbox open access? If not, please use generic term. Please reword 2.1 to show what is being done and how? e.g., set up tensor box by clicking xxxx or running the command xxxx
Response: We have made it clear. Yes, tensorbox is open access.

12. 2.1.1: How do you ensure?
Response: We have provided links to our github account.

13. Lines 111 onwards: Are these command lines or button clicks? Please be explicit.
Response: We have clarified these.

14. There is a 10-page limit for the Protocol, but there is a 2.75-page limit for filmable content. Please highlight 2.75 pages or less of the Protocol (including headings and spacing) that identifies the essential steps of the protocol for the video, i.e., the steps that should be visualized to tell the most cohesive story of the Protocol.
Response: We have highlighted the appropriate sections.

15. In the representative result section, please include the conclusion drawn from your protocol.
Response: We have provided this.

16. Please obtain explicit copyright permission to reuse any figures from a previous publication. Explicit permission can be expressed in the form of a letter from the editor or a link to the editorial policy that allows re-prints. Please upload this information as a .doc or .docx file to your Editorial Manager account. The Figure must be cited appropriately in the Figure Legend, i.e. "This figure has been modified from [citation]."
Response: The corresponding author of this manuscript is also the corresponding author of the other paper and we have obtained permission to adapt the figures. We have also checked with the other journal (Frontiers Systems Neuroscience) for copyright issues and confirmed that there are no issues.

17. As we are a methods journal, please revise the Discussion to explicitly cover the following in detail in 3-6 paragraphs with citations:
a) Critical steps within the protocol
b) Any modifications and troubleshooting of the technique
c) Any limitations of the technique
d) The significance with respect to existing methods
e) Any future applications of the technique
Response: We have modified accordingly.

Reviewers' comments:
Reviewer #1:
The authors present the unique observation that traditional behavior analyses can take a superficial approach to otherwise rich behavioral data sets. To address this shortcoming, the

authors propose the DeepBehavior toolbox. Step-by-step instructions to implement the DeepBehavior learning algorithms are included. While the authors cite their 2019 paper (Arac et al., 2019) which provides detailed examples of DeepBehavior's applications in both clinical and preclinical populations, the current paper would be strengthened by briefly emphasizing how DeepBehavior is distinct from typical behavioral analyses methods. This can be accomplished by succinctly comparing and contrasting applications and outputs of DeepBehavior toolbox with more typical behavioral analyses approaches. Additionally, implementation of the DeepBehavior toolbox requires a Linux operating system and working knowledge of Python. This is a limitation of the toolbox and should be stated early in the paper.

The authors provide detailed installation and implementation instructions and also comment on errors which may occur when using the toolbox, A notable strength of the toolbox proposed here, are the three different neural architectures and descriptions of when to use each. Importantly, DeepBehavior can be used to analyze both clinical and preclinical behavioral data and the paper would be strengthen by explicitly stating this early.

The authors should consider defining all acronyms the first time they are used to make the paper more user friendly for a diverse audience. However, the paper is well written and overall presents an exciting and novel methodology with a diverse range of applications and utility.
Response: We thank the reviewer for these comments. We have emphasized how this method is distinct from other typical behavior analysis methods, and also addressed the linux/python concern.

Reviewer #2:
Manuscript Summary:
The authors provide detailed step-by-step instructions with visualizations of our recent methodology, DeepBehavior, utilizing deep learning algorithms. The proposed method uses deep learning frameworks built with convolutional neural networks to rapidly process and analyze behavioral videos. The paper is of scientific and original nature, related to a step-by-step user guide for implementation of DeepBehavior: A deep learning toolbox for automated behavior analysis. These methods provide robust, automated, and precise ways to quantify behavioral tasks.

Minor Concerns: For a better clarification, please edit your paper as follows: Enlarge the Introduction with current results reported in the world and Europe, - References to expand the results of European authors registered in SCOPUS / WoS such as: Advanced Robotic Grasping System Using Deep Learning and Trends in Simulation and Planning of Manufacturing Companies. .Unify font in tables. Correct English grammar in this paper.
Please, edit the paper according to previous comments and after minor changes I recommend the paper to be published.
Response: We thank the reviewer for these comments. We have now revised the manuscript with recommended expansions, grammar corrections and appropriate references.

1 Alewife Center #200
Cambridge, MA 02140
tel. 617.945.9051
www.jove.com

ARTICLE AND VIDEO LICENSE AGREEMENT

Title of Article:	A step-by-step user guide for implementation of DeepBehavior: A deep learning toolbox for automated behavior analysis
Author(s):	Sanjay Shukla, Ahmet Arac

Item 1: The Author elects to have the Materials be made available (as described at http://www.jove.com/publish) via:

- [x] Standard Access
- [ ] Open Access

Item 2: Please select one of the following items:

- [x] The Author is **NOT** a United States government employee.
- [ ] The Author is a United States government employee and the Materials were prepared in the course of his or her duties as a United States government employee.
- [ ] The Author is a United States government employee but the Materials were NOT prepared in the course of his or her duties as a United States government employee.

ARTICLE AND VIDEO LICENSE AGREEMENT

1. **Defined Terms.** As used in this Article and Video License Agreement, the following terms shall have the following meanings: "**Agreement**" means this Article and Video License Agreement; "**Article**" means the article specified on the last page of this Agreement, including any associated materials such as texts, figures, tables, artwork, abstracts, or summaries contained therein; "**Author**" means the author who is a signatory to this Agreement; "**Collective Work**" means a work, such as a periodical issue, anthology or encyclopedia, in which the Materials in their entirety in unmodified form, along with a number of other contributions, constituting separate and independent works in themselves, are assembled into a collective whole; "**CRC License**" means the Creative Commons Attribution-Non Commercial-No Derivs 3.0 Unported Agreement, the terms and conditions of which can be found at: http://creativecommons.org/licenses/by-nc-nd/3.0/legalcode; "**Derivative Work**" means a work based upon the Materials or upon the Materials and other pre-existing works, such as a translation, musical arrangement, dramatization, fictionalization, motion picture version, sound recording, art reproduction, abridgment, condensation, or any other form in which the Materials may be recast, transformed, or adapted; "**Institution**" means the institution, listed on the last page of this Agreement, by which the Author was employed at the time of the creation of the Materials; "**JoVE**" means MyJove Corporation, a Massachusetts corporation and the publisher of The Journal of Visualized Experiments; "**Materials**" means the Article and / or the Video; "**Parties**" means the Author and JoVE; "**Video**" means any video(s) made by the Author, alone or in conjunction with any other parties, or by JoVE or its affiliates or agents, individually or in collaboration with the Author or any other parties, incorporating all or any portion of the Article, and in which the Author may or may not appear.

2. **Background.** The Author, who is the author of the Article, in order to ensure the dissemination and protection of the Article, desires to have the JoVE publish the Article and create and transmit videos based on the Article. In furtherance of such goals, the Parties desire to memorialize in this Agreement the respective rights of each Party in and to the Article and the Video.

3. **Grant of Rights in Article.** In consideration of JoVE agreeing to publish the Article, the Author hereby grants to JoVE, subject to **Sections 4** and **7** below, the exclusive, royalty-free, perpetual (for the full term of copyright in the Article, including any extensions thereto) license (a) to publish, reproduce, distribute, display and store the Article in all forms, formats and media whether now known or hereafter developed (including without limitation in print, digital and electronic form) throughout the world, (b) to translate the Article into other languages, create adaptations, summaries or extracts of the Article or other Derivative Works (including, without limitation, the Video) or Collective Works based on all or any portion of the Article and exercise all of the rights set forth in (a) above in such translations, adaptations, summaries, extracts, Derivative Works or Collective Works and(c) to license others to do any or all of the above. The foregoing rights may be exercised in all media and formats, whether now known or hereafter devised, and include the right to make such modifications as are technically necessary to exercise the rights in other media and formats. If the "Open Access" box has been checked in **Item 1** above, JoVE and the Author hereby grant to the public all such rights in the Article as provided in, but subject to all limitations and requirements set forth in, the CRC License.

4. **Retention of Rights in Article.** Notwithstanding the exclusive license granted to JoVE in **Section 3** above, the Author shall, with respect to the Article, retain the non-exclusive right to use all or part of the Article for the non-commercial purpose of giving lectures, presentations or teaching classes, and to post a copy of the Article on the Institution's website or the Author's personal website, in each case provided that a link to the Article on the JoVE website is provided and notice of JoVE's copyright in the Article is included. All non-copyright intellectual property rights in and to the Article, such as patent rights, shall remain with the Author.

5. **Grant of Rights in Video – Standard Access.** This **Section 5** applies if the "Standard Access" box has been checked in **Item 1** above or if no box has been checked in **Item 1** above. In consideration of JoVE agreeing to produce, display or otherwise assist with the Video, the Author hereby acknowledges and agrees that, Subject to **Section 7** below, JoVE is and shall be the sole and exclusive owner of all rights of any nature, including, without limitation, all copyrights, in and to the Video. To the extent that, by law, the Author is deemed, now or at any time in the future, to have any rights of any nature in or to the Video, the Author hereby disclaims all such rights and transfers all such rights to JoVE.

6. **Grant of Rights in Video – Open Access.** This **Section 6** applies only if the "Open Access" box has been checked in **Item 1** above. In consideration of JoVE agreeing to produce, display or otherwise assist with the Video, the Author hereby grants to JoVE, subject to **Section 7** below, the exclusive, royalty-free, perpetual (for the full term of copyright in the Article, including any extensions thereto) license (a) to publish, reproduce, distribute, display and store the Video in all forms, formats and media whether now known or hereafter developed (including without limitation in print, digital and electronic form) throughout the world, (b) to translate the Video into other languages, create adaptations, summaries or extracts of the Video or other Derivative Works or Collective Works based on all or any portion of the Video and exercise all of the rights set forth in (a) above in such translations, adaptations, summaries, extracts, Derivative Works or Collective Works and (c) to license others to do any or all of the above. The foregoing rights may be exercised in all media and formats, whether now known or hereafter devised, and include the right to make such modifications as are technically necessary to exercise the rights in other media and formats. For any Video to which this **Section 6** is applicable, JoVE and the Author hereby grant to the public all such rights in the Video as provided in, but subject to all limitations and requirements set forth in, the CRC License.

7. **Government Employees.** If the Author is a United States government employee and the Article was prepared in the course of his or her duties as a United States government employee, as indicated in **Item 2** above, and any of the licenses or grants granted by the Author hereunder exceed the scope of the 17 U.S.C. 403, then the rights granted hereunder shall be limited to the maximum rights permitted under such statute. In such case, all provisions contained herein that are not in conflict with such statute shall remain in full force and effect, and all provisions contained herein that do so conflict shall be deemed to be amended so as to provide to JoVE the maximum rights permissible within such statute.

8. **Protection of the Work.** The Author(s) authorize JoVE to take steps in the Author(s) name and on their behalf if JoVE believes some third party could be infringing or might infringe the copyright of either the Author's Article and/or Video.

9. **Likeness, Privacy, Personality.** The Author hereby grants JoVE the right to use the Author's name, voice, likeness, picture, photograph, image, biography and performance in any way, commercial or otherwise, in connection with the Materials and the sale, promotion and distribution thereof. The Author hereby waives any and all rights he or she may have, relating to his or her appearance in the Video or otherwise relating to the Materials, under all applicable privacy, likeness, personality or similar laws.

10. **Author Warranties.** The Author represents and warrants that the Article is original, that it has not been published, that the copyright interest is owned by the Author (or, if more than one author is listed at the beginning of this Agreement, by such authors collectively) and has not been assigned, licensed, or otherwise transferred to any other party. The Author represents and warrants that the author(s) listed at the top of this Agreement are the only authors of the Materials. If more than one author is listed at the top of this Agreement and if any such author has not entered into a separate Article and Video License Agreement with JoVE relating to the Materials, the Author represents and warrants that the Author has been authorized by each of the other such authors to execute this Agreement on his or her behalf and to bind him or her with respect to the terms of this Agreement as if each of them had been a party hereto as an Author. The Author warrants that the use, reproduction, distribution, public or private performance or display, and/or modification of all or any portion of the Materials does not and will not violate, infringe and/or misappropriate the patent, trademark, intellectual property or other rights of any third party. The Author represents and warrants that it has and will continue to comply with all government, institutional and other regulations, including, without limitation all institutional, laboratory, hospital, ethical, human and animal treatment, privacy, and all other rules, regulations, laws, procedures or guidelines, applicable to the Materials, and that all research involving human and animal subjects has been approved by the Author's relevant institutional review board.

11. **JoVE Discretion.** If the Author requests the assistance of JoVE in producing the Video in the Author's facility, the Author shall ensure that the presence of JoVE employees, agents or independent contractors is in accordance with the relevant regulations of the Author's institution. If more than one author is listed at the beginning of this Agreement, JoVE may, in its sole

discretion, elect not take any action with respect to the Article until such time as it has received complete, executed Article and Video License Agreements from each such author. JoVE reserves the right, in its absolute and sole discretion and without giving any reason therefore, to accept or decline any work submitted to JoVE. JoVE and its employees, agents and independent contractors shall have full, unfettered access to the facilities of the Author or of the Author's institution as necessary to make the Video, whether actually published or not. JoVE has sole discretion as to the method of making and publishing the Materials, including, without limitation, to all decisions regarding editing, lighting, filming, timing of publication, if any, length, quality, content and the like.

12. **Indemnification.** The Author agrees to indemnify JoVE and/or its successors and assigns from and against any and all claims, costs, and expenses, including attorney's fees, arising out of any breach of any warranty or other representations contained herein. The Author further agrees to indemnify and hold harmless JoVE from and against any and all claims, costs, and expenses, including attorney's fees, resulting from the breach by the Author of any representation or warranty contained herein or from allegations or instances of violation of intellectual property rights, damage to the Author's or the Author's institution's facilities, fraud, libel, defamation, research, equipment, experiments, property damage, personal injury, violations of institutional, laboratory, hospital, ethical, human and animal treatment, privacy or other rules, regulations, laws, procedures or guidelines, liabilities and other losses or damages related in any way to the submission of work to JoVE, making of videos by JoVE, or publication in JoVE or elsewhere by JoVE. The Author shall be responsible for, and shall hold JoVE harmless from, damages caused by lack of sterilization, lack of cleanliness or by contamination due to the making of a video by JoVE its employees, agents or independent contractors. All sterilization, cleanliness or decontamination procedures shall be solely the responsibility of the Author and shall be undertaken at the Author's expense. All indemnifications provided herein shall include JoVE's attorney's fees and costs related to said losses or damages. Such indemnification and holding harmless shall include such losses or damages incurred by, or in connection with, acts or omissions of JoVE, its employees, agents or independent contractors.

13. **Fees.** To cover the cost incurred for publication, JoVE must receive payment before production and publication of the Materials. Payment is due in 21 days of invoice. Should the Materials not be published due to an editorial or production decision, these funds will be returned to the Author. Withdrawal by the Author of any submitted Materials after final peer review approval will result in a US$1,200 fee to cover pre-production expenses incurred by JoVE. If payment is not received by the completion of filming, production and publication of the Materials will be suspended until payment is received.

14. **Transfer, Governing Law.** This Agreement may be assigned by JoVE and shall inure to the benefits of any of JoVE's successors and assignees. This Agreement shall be governed and construed by the internal laws of the Commonwealth of Massachusetts without giving effect to any conflict of law provision thereunder. This Agreement may be executed in counterparts, each of which shall be deemed an original, but all of which together shall be deemed to me one and the same agreement. A signed copy of this Agreement delivered by facsimile, e-mail or other means of electronic transmission shall be deemed to have the same legal effect as delivery of an original signed copy of this Agreement.

A signed copy of this document must be sent with all new submissions. Only one Agreement is required per submission.

CORRESPONDING AUTHOR

Name:	Ahmet Arac
Department:	Neurology
Institution:	UCLA
Title:	Assistant Professor
Signature:	Ahmet Arac
Date:	09/03/2019

Please submit a **signed** and **dated** copy of this license by one of the following three methods:

1. Upload an electronic version on the JoVE submission site
2. Fax the document to +1.866.381.2236
3. Mail the document to JoVE / Attn: JoVE Editorial / 1 Alewife Center #200 / Cambridge, MA 02140

Signature Certificate

Document Ref.: TBG8T-HN22Y-CMDZ6-CDCDF

Document signed by:

Ahmet Arac

Verified E-mail:
aarac@mednet.ucla.edu

IP: 76.175.67.167 | Date: 04 Sep 2019 06:47:11 UTC

Ahmet Arac

Document completed by all parties on:
04 Sep 2019 06:47:11 UTC