EAI Endorsed Transactions
on Bioengineering and Bioinformatics

Research Article EAI.EU

Estimating animal pose using deep learning: a trained deep learning model outperforms morphological analysis

S. Lee[1,*], J. Banzon[1], K. Le[2] and D. H. Kim[2]

[1]Marshall University, Huntington, WV, USA
[2]Kennesaw State University, Marietta, GA, USA

Abstract

INTRODUCTION: Analyzing animal behavior helps researchers understand their decision-making process and helper tools are rapidly becoming an indispensable part of many interdisciplinary studies. However, researchers are often challenged to estimate animal pose because of the limitation of the tools and its vulnerability to a specific environment. Over the years, deep learning has been introduced as an alternative solution to overcome these challenges.
OBJECTIVES: This study investigates how deep learning models can be applied for the accurate prediction of animal behavior, comparing with traditional morphological analysis based on image pixels.
METHODS: Transparent Omnidirectional Locomotion Compensator (TOLC), a tracking device, is used to record videos with a wide range of animal behavior. Recorded videos contain two insects: a walking red imported fire ant (*Solenopsis invicta*) and a walking fruit fly (*Drosophila melanogaster*). Body parts such as the head, legs, and thorax, are estimated by using an open-source deep-learning toolbox. A deep learning model, ResNet-50, is trained to predict the body parts of the fire ant and the fruit fly respectively. 500 image frames for each insect were annotated by humans and then compared with the predictions of the deep learning model as well as the points generated from the morphological analysis.
RESULTS: The experimental results show that the average distance between the deep learning-predicted centroids and the human-annotated centroids is 2.54, while the average distance between the morphological analysis-generated centroids and the human-annotated centroids is 6.41 over the 500 frames of the fire ant. For the fruit fly, the average distance of the centroids between the deep learning- predicted and the human-annotated is 2.43, while the average distance of the centroids between the morphological analysis-generated and the human-annotated is 5.06 over the 477 image frames.
CONCLUSION: In this paper, we demonstrate that the deep learning model outperforms traditional morphological analysis in terms of estimating animal pose in a series of video frames.

Keywords: animal pose estimation, deep learning, and morphological analysis.

Received on 21 February 2022, accepted on 20 April 2022, published on 22 April 2022

Copyright © 2022 S. Lee *et al.*, licensed to EAI. This is an open-access article distributed under the terms of the Creative Commons Attribution license, which permits unlimited use, distribution and reproduction in any medium so long as the original work is properly cited.

doi: 10.4108/eai.22-4-2022.173951

1. Introduction

Understanding the relationship between nervous system function and behavior has been a major goal in the neuroscience research [2]. Real et al. [3] presented a circuit inference framework that represents neurons and their connections to facilitate the understanding of neural circuits, and Deisseroth [4] emphasized the significance of the function of nervous systems. While a variety of studies has been focused on the structure of neurons and function, the analysis of animal behavior has been moved slowly because of the lack of understanding about theoretical and experimental research of behavior [5].

*Corresponding author. Email: leesan@marshall.edu

Animal behavior is an important part of neuroscience study and has been treated in many fields, such as genetics, psychology, and ethology [6]. To understand animal behavior, high-quality video recordings have been used to capture the motion signal of behaving animals. However, researchers have been suffered from the lack of an efficient method for the animal behavior analysis in the captured video recordings; the analysis takes too much time to complete all tasks for each frame in the captured video recordings. Moreover, it may be highly subjective judging whether the animal behavior changes in certain situations.

Advances in automatic technologies have enabled animal scientists or researchers to better study the behavior of animals, providing the automated analysis of images [7, 8]. In bioimage analysis, automated image analysis using a convolutional neural network (CNN) has been a popular method for studying digital images. As a class of neural network model, the CNN-based models, such as CNN [9], Recurrent CNN-based models [10], encoder-decoder-based models, generative adversarial networks [11], and CNN-based active learning models [12] have been widely studied and successfully applied to the real-world problems [13-15]. In particular, Apte et al. presented that a deep learning-based image segmentation can be involved in an open-source library [16] and Stern et al. presented a high-throughput analysis of animal behavior by using CNN [17]. Further, some studies show CNN- based models can improve the fundamental image analysis operations [18, 19]. However, these methods are still time-consuming and require a heavy workforce for training and testing on the captured video recordings.

Animal pose estimation has been received much attention by reducing the time-consuming effort to label. Newell et al. [20] introduced a CNN architecture called stacked hourglass for human pose estimation and Insafutdinov et al. [21] firstly presented an articulated multi-person pose estimation by using the most popular deep learning model, ResNet. Mathis et al. [1] advanced Insafutdinov's work. However, these methods have been focused on deep learning models and didn't demonstrate their models by comparing them with morphological analysis.

In this paper, we investigate how CNN-based models can achieve accurate animal pose predictions by using a newly developed open-source software toolbox called DeepLabCut [1, 22], comparing with traditional morphological image analysis. Raw images of a fruit fly (*Drosophila melanogaster*) and *fire ant* (*Solenopsis invicta*) are recorded by a motion-tracking device called transparent omnidirectional locomotion compensator (TOLC) [23] separately. Each video recording created by TOLC is trained by deep learning models and the trained deep learning models are used to predict animal poses. The animal poses predicted by the trained deep learning models are compared with the poses estimated by morphological images analysis.

The rest of the paper is organized as follows. In Section 2, we describe our TOLC device. In Section 3, we introduce the DeepLabCut open-software tool. To provide a better understanding of the raw image data and the methods used in this paper, Section 4 shows detailed information about the measurement, raw images, and model configuration. In Section 5, we present our experimental results. Finally, Section 6 concludes the presented work.

2. Materials and Method

2.1. Tracking device

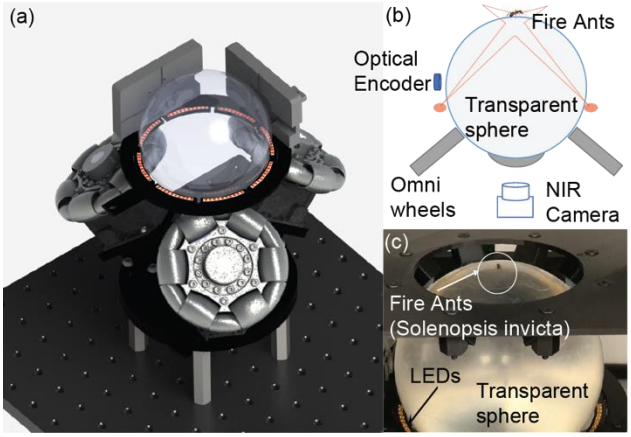

Figure 1. (a) Omnidirectional Motion Compensation System. (b) The schematic drawing of main components. (c). closed-view of fire ants on the tracking device

Tracking devices have been being developed extensively by animal scientists or researchers because tracking behavioral data help to capture footage from an animal's movement and provides invaluable contributions to many disciplines such as genetics, psychology, and ethology [24, 25]. This tracking device records the behavior of walking insects, such as a fruit fly, a cricket, and ants, by compensating their motion using a transparent sphere. However, the behavior observation system has been having difficulty obtaining all behavioral data because of its limited field of view. To address this issue, much attention has been paid to the development of a finite behavior chamber to observe the behavior in a freely walking insect by restricting the behavior of the animal.

Meanwhile, an insect was tethered on the fixture and walked on top of the air-floated ball in the finite behavior chamber. However, the tethered method makes an insect move its legs only to rotate a ball while its body is immobilized in the fixture. Although this method allows infinite space navigation with a virtual reality system, it causes a serious problem that damages the animal during the tethering process and makes it difficult for a long-term behavior study of an insect. We extended our paper [26] to include a walking red imported fire ant (*Solenopsis invicta*) without tethering but using the feedback control of the TOLC.

The TOLC [23] is a tracking device that captures an insect's movement by compensating its motion with the help of a transparent sphere. This TOLC device, shown in Figure 1, detects the motion of the untethered walking insect and brings it back to the desired location by rotating a sphere to counteract the walking motion. The imaging system underneath the transparent sphere captures an image of the freely walking insect at 200 Hz with the illumination of 850 nm near-infrared (NIR) light-emitting diodes (LEDs) to minimize the visual distraction of the insect. Once the position is detected, three omnidirectional wheels attached to the servomotor rotate the sphere in the desired direction to cancel out error between the current and target positions. We implemented the proportional integral derivative(PID) control methods to compute the control input [23]. The rotation is measured by two optical encoders on the side of the sphere, which is used to estimate the travel trajectory of the walking insect.

2.2. Pose estimation

Pose estimation is a machine learning task that estimates the object position from a series of image frames taken from a video recording at a particular frame rate [27]. The object position has been detected, associated, and tracked by a model reference frame obtained at a previous time, but the accuracy of pose estimation has been limited by high computational resources. Today, advances in image analysis methods based on convolutional neural networks (CNNs) promote fast and accurate object pose inference by pre-trained models for pose detection and pose tracking.

DeepLabCut is an image analysis tool that provides capabilities for tracking behaviors of various objects over state-of-the-art deep learning models. This deep learning-equipped image analysis tool helps users to easily label the subject pose and train the labels for the object pose inference. DeepLabCut provides users steps to successfully predict the object pose. The steps follow: creating a project, extracting frames, labeling object pose, training labels, and evaluating deep learning models. DeepLabCut also provides the ability to retrain evaluated deep learning models by extracting outlier frames or updating labeled frames.

DeepLabCut utilizes deep residual networks such as ResNet-50 and ResNet-101 which are ImageNet pre-trained to estimate human body parts with high performance. ResNet introduced in [28] is a special type of neural network that has received much attention in many image vision communities because a large number of layers on neural networks avoid training errors. As the state-of-the-art neural network architectures, such as VGG network [29], GoogleNet [30], and SqueezeNet [31], have been suffered in the training performance of large layered networks, ResNet models provided a solution to the problem by fitting the stacked layers to a residual mapping.

Moreover, DeepLabCut offers several benefits over similar available tools such as DataJoint, Kinematic, and Openephys, for feature tracking. DeepLabCut guides the experimenter using an interface that provides a step-by-step procedure from labeling to training, minimizes the cost of manual behavior analysis and can achieve human-level accuracy with only a small amount of labeled images, can be easily adapted to analyze behaviors across species, and is open-source and free to use.

The procedure of DeepLabCut is established by several steps: extracting the region of interest (ROI), manually localizing body parts, training a deep neural network (DNN) architecture, and predicting the locations of the body parts from new videos. Particularly, the DNN architecture uses ResNet-50 to predict the location of the body part by updating a distinct readout layer and network weights. The architecture of DeepLabCut with ResNet-50 is shown in Figure 2.

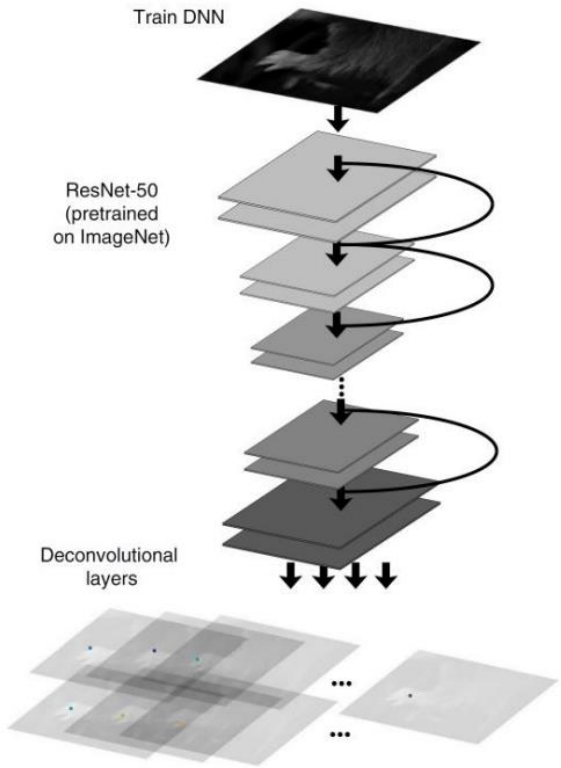

Figure 2. DeepLabCut ResNet-50 architecture [1].

In this paper, we use ResNet-50 built in the DeepLabCut and compare it with traditional morphological analysis. Especially, we investigate how the centroid of an object is accurately predicted by the deep learning model. The results of the deep learning model will be compared with the centroids measured by using morphological image processing in a sequence of the image frames.

2.3. Datasets

We created two datasets for *Drosophila melanogaster* and *fire ant*. For the experiment of estimating *Drosophila*

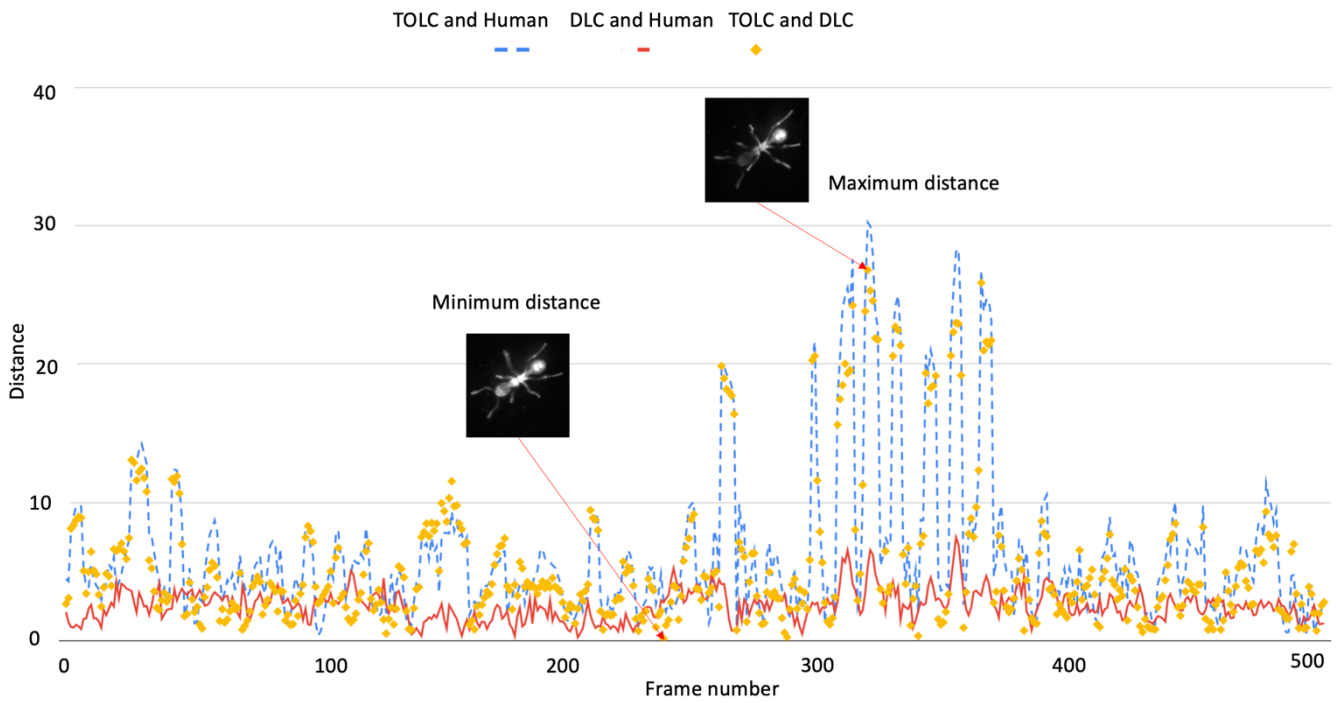

Figure 3. Experimental results of the distance comparisons of deep learning-predicted (DLC), morphological analysis-generated (TOLC), and human-annotated (Human) on the fire ant.

melanogaster pose, wild-type Canton S was obtained from the *Drosophila* stock center (Bloomington, IN) to be raised with a standard food (Genesee Scientific, 66-117) at 25°C in a 12-hour dark-light cycle. Fruit flies between twenty and 40 days old were used for the first experiment in the chamber under CO_2 anesthesia. The fire ant workers used for experiments were collected from the campus of Kennesaw State University (Georgia, USA). Fire ants were housed in a plastic chamber with a nest tube, and they were provided with sugar water and held at standard room temperature (~22°C). For an initial 5 minutes, both the fruit fly and the fire ant were left to habituate before the first experiment. The experiment is conducted in the dark chamber to eliminate unpredicted external stimuli.

To record two animals, the near-infrared (NIR) camera (FLIR, CM3-U3-31S4M-CS) with 850 nm light-emitting diodes (Osram, SFH 4655-Z) equipped in TOLC was used to capture the animals' motion. Videos were recorded and their metadata were formatted as a binary file with 640x376 sized images and 720x540 sized images for a fruit fly and a fire ant respectively.

Experiments were performed on one Ubuntu server with 40 cores, 196 GB memory, and two Quadro P4000 GPUs. ResNet-50 was adopted to train centroids of a fruit fly and a fire ant. We obtained 1,000 fruit fly images and 1,000 fire ant images from the videos splitting them into two 500 images for training and testing. The number of training iterations was 15,000. A cross-entropy loss function with a learning rate of 0.02 was used for training. The cross-entropy loss function represents the L1 distance between a predicted region and an actual region. The loss function is defined as below:

$$(\hat{y}, y) = -(y\log\hat{y} + (1-y)\log(1-\hat{y})) \qquad (1)$$

, where $\hat{y}$ is a predicted region in pixel and y is an actual region in pixel.

The location of the fruit fly's centroid was measured by a morphological filter from the TOLC device. The morphological filter computes the centroids in 2-dimensional space with x-axis and y-axis. The geometric center (tc_x, tc_y) of the fruit fly's area was compared with the location (dc_x, dc_y) of the centroid predicted by the ResNet-50 in the DeepLabCut. Likewise, we obtained the centroid location of the fire ant by the morphological filter and the deep learning prediction, comparing them with each other.

Two experiments were conducted for the fruit fly and the fire ant respectively. To verify the effectiveness of the deep learning model, we compared the (dc_x, dc_y) with the location (hc_x, hc_y) of the centroid annotated by two students (one graduate student and one undergraduate student), human-annotated. In the labeling process, the centroids in 50 image frames were labeled for the training of both animals. For the testing, 500 image frames were labeled using DeepLabCut's graphical user interface (GUI).

3. Experiment Results

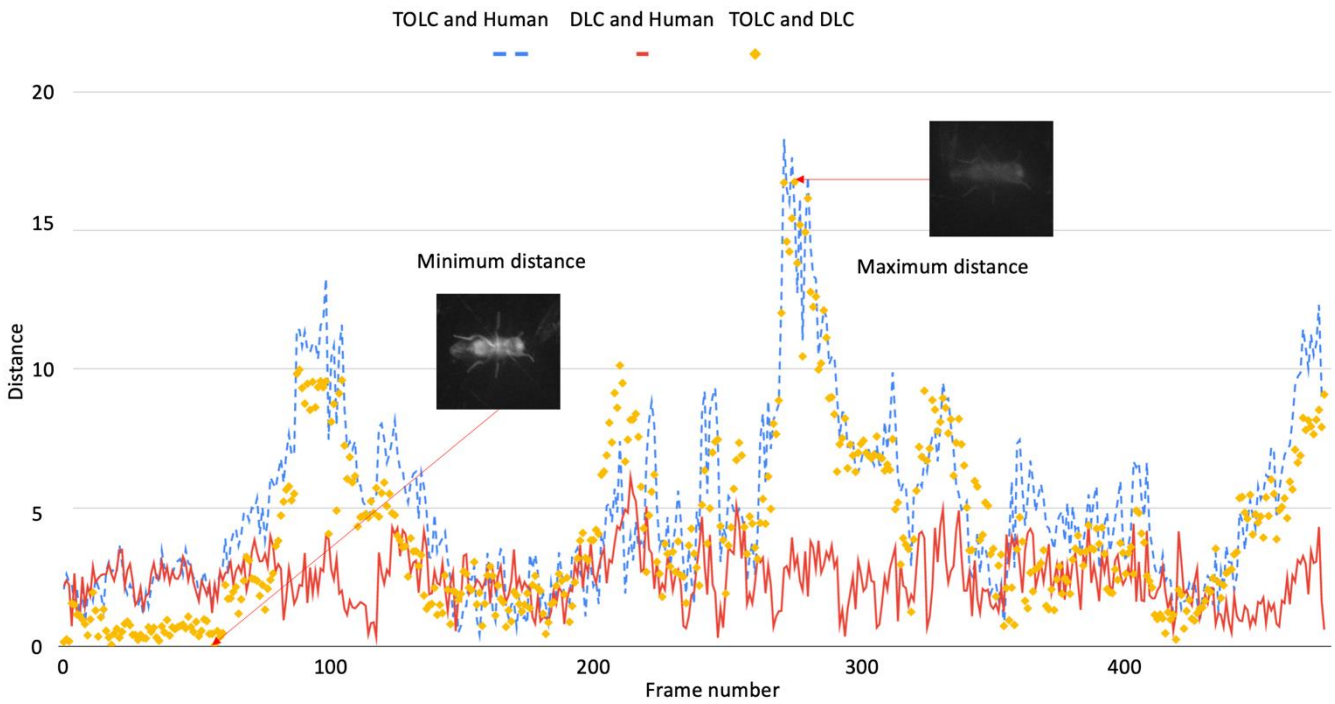

Figure 4. Experimental results of the distance comparisons of deep learning-predicted (DLC), morphological analysis-generated (TOLC), and human-annotated (Human) on the fruit fly.

In this section, we demonstrate the effectiveness of the deep learning model based on the results of the experiments comparing three centroids: (tc_x, tc_y) calculated by the morphological analysis, (dc_x, dc_y) predicted by the deep learning model, and (hc_x, hc_y) annotated by humans.

The Euclidean distance was chosen to compare the distance between two centroids to find the dissimilarity of the centroid location. The Euclidean distances $ED(d,h)$, $ED(d,t)$, and $ED(t,h)$ used in the experiment are defined as below:

$$ED(d,h) = \sqrt{(dc_x - hc_x)^2 + (dc_y - hc_y)^2} \quad (2)$$

$$ED(d,t) = \sqrt{(dc_x - tc_x)^2 + (dc_y - tc_y)^2} \quad (3)$$

$$ED(t,h) = \sqrt{(tc_x - hc_x)^2 + (tc_y - hc_y)^2} \quad (4)$$

The distance between deep learning-predicted centroids and human-annotated centroids is defined as $ED(d,h)$, while the distance between deep learning-predicted centroids and morphological analysis-generated centroids is defined as. The distance between morphological analysis-generated centroids and human-annotated centroids is defined as $ED(t,h)$.

The experiment results of Euclidean distances: $ED(d,h)$, $ED(d,t)$, and $ED(t,h)$ on the testing datasets for the fire ant are shown in Figure 3. In the figure, the x-axis represents frame numbers recorded by the TOLC device, while the y-axis represents the distances measured by Euclidean distance. The solid red line representing $ED(d,h)$ indicates that the centroid distances are likely to be maintained as the frame numbers increase, while the dotted blue line representing $ED(t,h)$ shows irregular distances as the frame numbers increase, comparing with the solid red line.

The overall experiment results of $ED(d,h)$ $ED(d,h)$, $ED(d,t)$, and $ED(t,h)$ on the testing datasets for the fire ant show that the deep learning-predicted centroids are much closer to the human-annotated centroids than the morphological analysis-generated centroids. These results demonstrate that the error distance of the centroids of the morphological analysis generated is much larger than the error distance of the centroids of the deep learning-predicted.

The Euclidean distances for $ED(d,h)$, $ED(d,t)$, and $ED(t,h)$ on the testing datasets for the fruit fly are shown in Figure 4. Similarly, the x-axis represents frame numbers recorded by the TOLC device, while the y-axis represents the distances measured by Euclidean distance. Unlike that the morphological analysis-generated centroids are fully obtained (500 frames) on the testing datasets for the fire ant, 23 morphological analysis-generated centroids were missing on the testing datasets for the fruit fly. Therefore, we performed the experiment on 477 frames of the fruit fly. These new unseen datasets were used to determine if the deep learning model was trained effectively or not. As a result, we found that the sold red line representing $ED(d,h)$ keeps maintaining the centroid distances between deep-learning predicted and human-annotated,

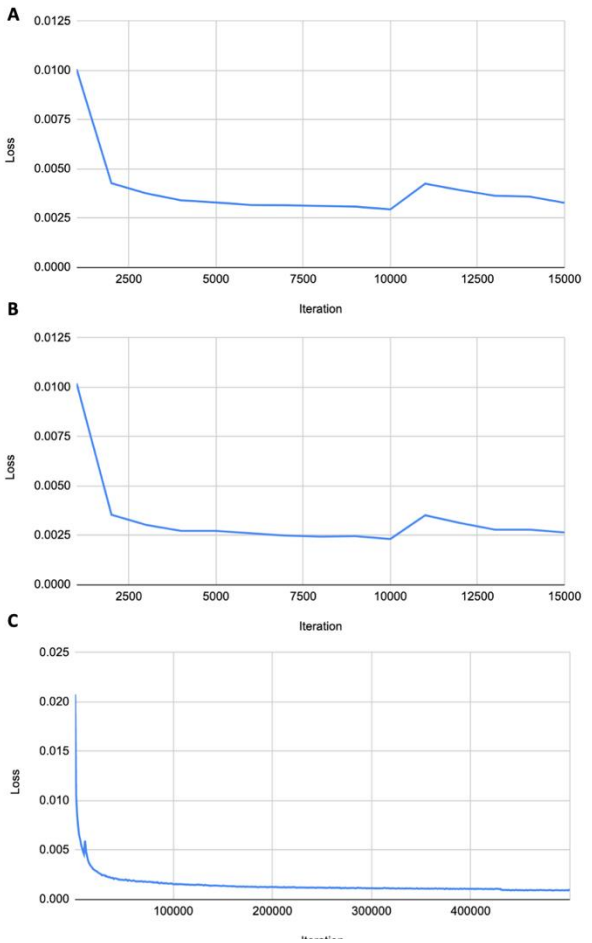

Figure 5. Loss over the iteration on training datasets. A. Loss on the fire ant centroids B. Loss on the fruit fly. C. Loss on the fire ant seven locations.

showing that the deep learning-predicted centroids are very similar to the human-annotated centroid. The dotted blue line representing $ED(t,h)$ provides a different result showing that the centroid distances between morphological analysis-generated and human-annotated are irregular as the frame numbers increase.

The cross-entropies $L(\hat{y},y)$ for training at 15,000 iterations were 0.0062 and 0.0062 with a 0.02 learning rate for the fruit fly and the fire ant respectively. The average distances between $ED(d,h)$, $ED(d,t)$, and $ED(t,h)$ are shown in Table 1.

Table 1. The average distances between $ED(d,\ h)$, $ED(d,\ t)$, and $ED(t,\ h)$.

	Average distance	
	Fruit fly	Fire ant
$\boldsymbol{ED(d,h)}$	**2.43**	**2.54**
$\boldsymbol{ED(t,h)}$	5.06	6.41
$\boldsymbol{ED(d,t)}$	4.14	5.49

An additional experiment to further validate our experiment was performed on other locations such as head, front head, front right leg, front left leg, middle right leg, middle left leg, rear right leg, and rear left leg. The number of training iterations was 500,000. A cross-entropy loss function representing the L1 distance between a predicted region and an actual region was used for training and the learning rate is 0.02. Training loss for each experiment is shown in Figure 5. While the experiment on centroids for both the fire ant and the fruity fly are conducted over the 15,000 iterations, we used 500,000 iterations for the seven locations. Three experiments indicate that the loss drops very quickly until around 2,300-th iteration and becomes constant as the iteration increases, demonstrating the effectiveness of the deep learning model on the experiments.

The visual representation of the locations on the fire ant is shown in Figure 6. Figure 6 represents fire ant NIR images with different times. The seven locations were colored as purple, blue, yellow, light blue, orange, light green, and red respectively.

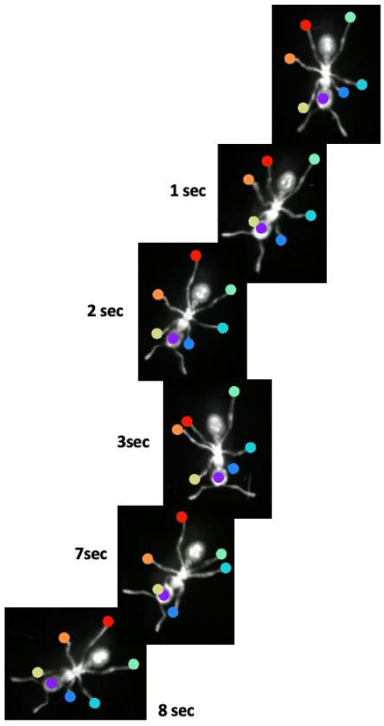

Figure 6. Visual representation of the DeepLabCut-predicted locations for head, front right leg, front left leg, middle right leg, middle left leg, rear right leg, and rear left leg.

The trajectories for the seven locations of the fire ant are shown in Figure 7. Figure 7A shows the fire ant's trajectories for the front legs. The upper figure indicates that both the front right leg and the front left leg have similar patterns across the 2-dimensional space. The bottom figure displays x and y-position over the image frames as time moves on. Likewise, the trajectories for both the rear legs and the middle legs produce similar results to

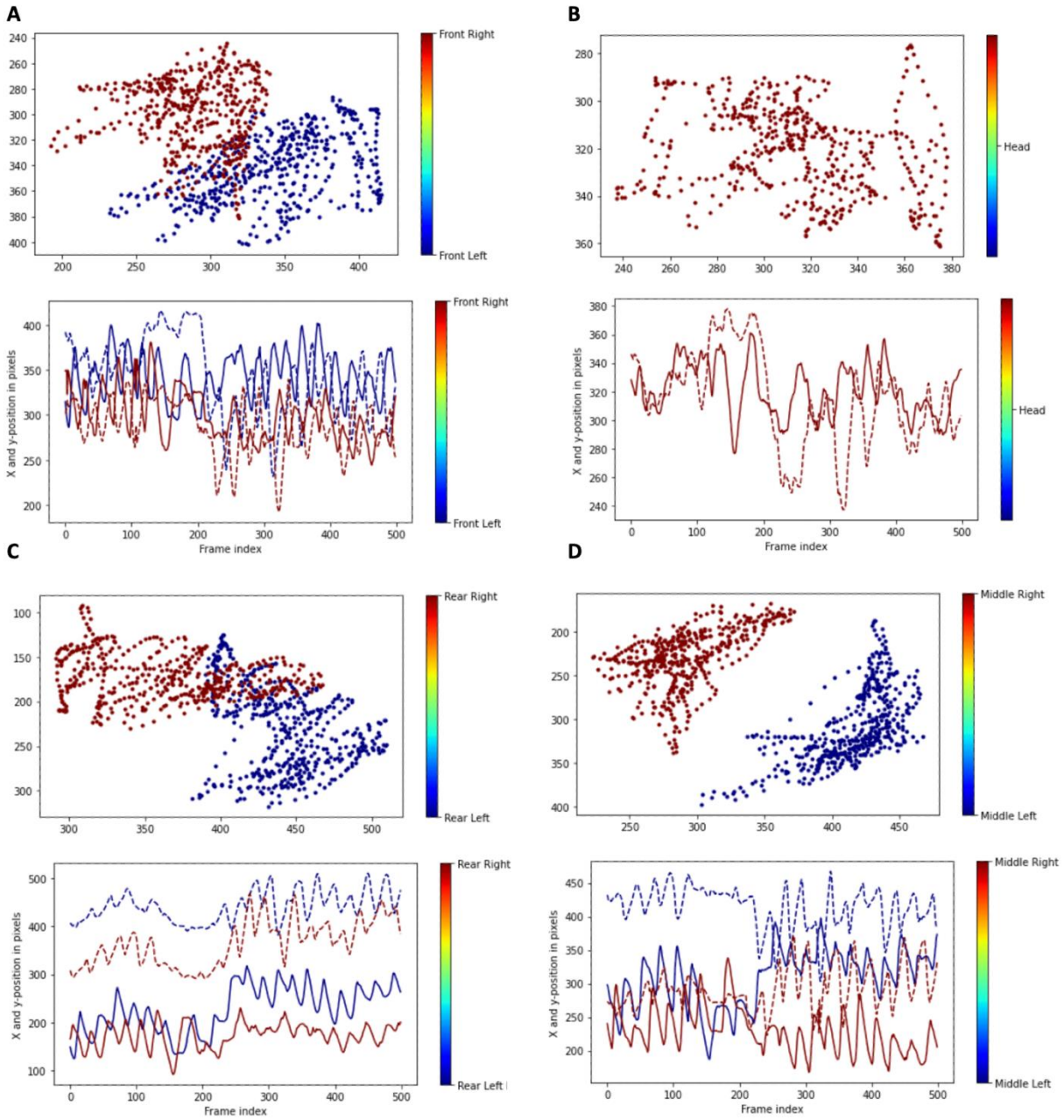

Figure 7. Ant trajectories for the seven locations. A. Front right leg and front left leg. B. Head. C. Rear right leg and rear left leg. D. Middle right leg and middle left leg.

the front legs. The trajectory information can be useful when researchers investigate animal behavior over the time domain. An example of the utilization of trajectory information is shown in Figure 8.

4. Discussion

Our study demonstrates that DeepLabCut can be applied to different species of animals and is beneficial to auto-tracking with a minimum of effort to identify the pattern of behavior exhibited by different species of animals. Our experimental results provide guidelines on how the predicted location information (i.e., centroids, head, and the six legs) can be combined. We plan to perform extensive experiments with more annotations that will avoid an intra-observer variation or learning bias. Moreover, we plan to study the conditioned behavior (i.e., positive phototactic movement) in *Drosophila* by tracking the leg movements. The phototactic response is the cellular behavior of moving directionally in response to the light source. While the negative phototactic response is related to moving away from the light, the positive phototactic response is related to moving toward the light. P. Pun et al. experimented on *Drosophila melanogaster*'s positive

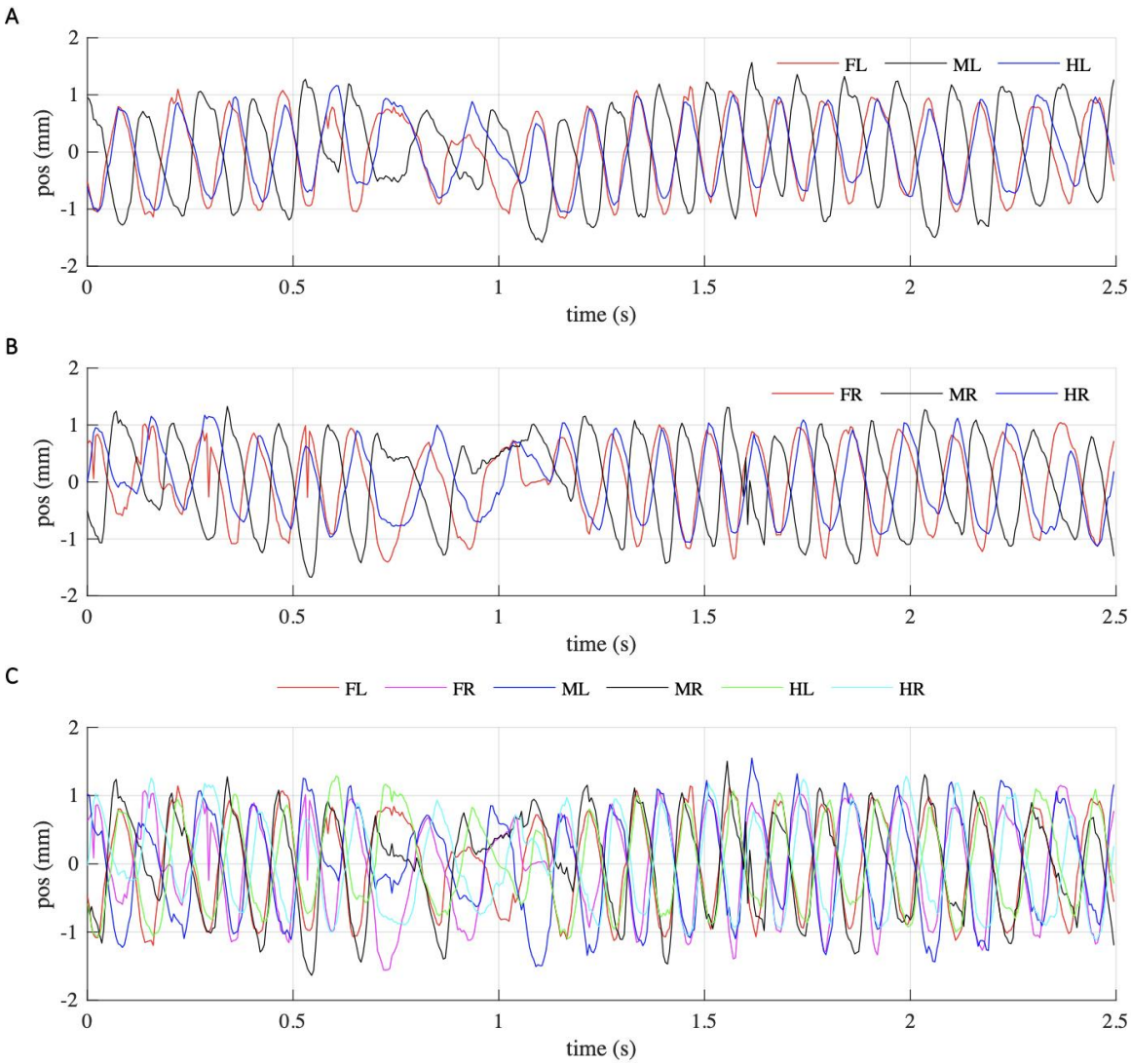

Figure 8. Comparison of leg positions over time obtained by ant trajectories. A. Front left legs, middle left legs, and hind left legs. B. Front right legs, middle right legs, and hind right legs. C. All legs Rear right leg and rear left leg. D. Middle right leg and middle left leg.

phototactic movement by using the TOLC device and demonstrated the ability of the TOLC device to detect the fruit fly's positive phototactic movement. This experiment would encourage us to use the sequence of DeepLabCut predicted leg movements with the heading angle measured by the TOLC device.

5. Conclusion

In this paper, we performed an estimation of the animal pose in a sequence of image frames recorded by the TOLC device. Poses of two insects, a fruit fly, and a fire ant, were estimated to explain patterns at the location of the body parts. Morphological image analysis was performed to estimate the centroid of a body area, comparing it with the centroid predicted by a deep learning model. While the morphological process highly depends on a pixel-wise mask, the deep learning process creates a model that enables us to perform a markerless pose estimation to locate the centroid of a body area. In addition to the experiment, we studied the patterns of the insets' movement by tracing the location of the head and six legs. A total of 2,000 image frames was used for the three experiments and 1,000 image frames were annotated for the centroids for both the fire ant and the fruit fly respectively. The experiment results showed that the overall distance between the deep learning-predicted centroids and the human-annotated centroids is less than the overall distance between the morphological analysis-generated centroids and the human-annotated centroids.

References

1 Mathis, A., Mamidanna, P., Cury, K.M., Abe, T., Murthy, V.N., Mathis, M.W., and Bethge, M.: 'DeepLabCut: markerless pose estimation of user-defined body parts with deep learning', Nature neuroscience, 2018, 21, (9), pp. 1281-1289
2 Pereira, T.D., Aldarondo, D.E., Willmore, L., Kislin, M., Wang, S.S.-H., Murthy, M., and Shaevitz, J.W.: 'Fast animal pose estimation using deep neural networks', Nature methods, 2019, 16, (1), pp. 117-125
3 Real, E., Asari, H., Gollisch, T., and Meister, M.: 'Neural circuit inference from function to structure', Current Biology, 2017, 27, (2), pp. 189-198
4 Deisseroth, K.: 'Optogenetics: 10 years of microbial opsins in neuroscience', Nature neuroscience, 2015, 18, (9), pp. 1213-1225
5 Krakauer, J.W., Ghazanfar, A.A., Gomez-Marin, A., MacIver, M.A., and Poeppel, D.: 'Neuroscience needs behavior: correcting a reductionist bias', Neuron, 2017, 93, (3), pp. 480-490
6 McFarland, D.: 'Animal behaviour: psychobiology, ethology, and evolution' (John Wiley & Sons, 1993. 1993)
7 LeCun, Y., Bengio, Y., and Hinton, G.: 'Deep learning', nature, 2015, 521, (7553), pp. 436-444
8 Buls, N., Watté, N., Nieboer, K., Ilsen, B., and de Mey, J.: 'Performance of an artificial intelligence tool with real-time clinical workflow integration–Detection of intracranial hemorrhage and pulmonary embolism', Physica Medica: European Journal of Medical Physics, 2021, 83, pp. 154-160
9 Yamashita, R., Nishio, M., Do, R.K.G., and Togashi, K.: 'Convolutional neural networks: an overview and application in radiology', Insights into imaging, 2018, 9, (4), pp. 611-629
10 Liang, M., and Hu, X.: 'Recurrent convolutional neural network for object recognition', in Editor (Ed.)^(Eds.): 'Book Recurrent convolutional neural network for object recognition' (2015, edn.), pp. 3367-3375
11 Wolterink, J.M., Leiner, T., Viergever, M.A., and Išgum, I.: 'Generative adversarial networks for noise reduction in low-dose CT', IEEE transactions on medical imaging, 2017, 36, (12), pp. 2536-2545
12 Lee, S., Amgad, M., Mobadersany, P., McCormick, M., Pollack, B.P., Elfandy, H., Hussein, H., Gutman, D.A., and Cooper, L.A.: 'Interactive classification of whole-slide imaging data for cancer researchers', Cancer Research, 2021, 81, (4), pp. 1171-1177
13 Kenner, B., Chari, S.T., Kelsen, D., Klimstra, D.S., Pandol, S.J., Rosenthal, M., Rustgi, A.K., Taylor, J.A., Yala, A., and Abul-Husn, N.: 'Artificial Intelligence and Early Detection of Pancreatic Cancer: 2020 Summative Review', Pancreas, 2021, 50, (3), pp. 251
14 Kumar, R., Khan, F.U., Sharma, A., Aziz, I., and Poddar, N.K.: 'Recent Applications of Artificial Intelligence in the Detection of Gastrointestinal, Hepatic and Pancreatic Diseases', Current Medicinal Chemistry, 2022, 29, (1), pp. 66-85
15 Loeffler, C.M.L., Bruechle, N.O., Jung, M., Seillier, L., Rose, M., Laleh, N.G., Knuechel, R., Brinker, T.J., Trautwein, C., and Gaisa, N.T.: 'Artificial Intelligence–based Detection of FGFR3 Mutational Status Directly from Routine Histology in Bladder Cancer: A Possible Preselection for Molecular Testing?', European Urology Focus, 2021
16 Apte, A.P., Iyer, A., Thor, M., Pandya, R., Haq, R., Jiang, J., LoCastro, E., Shukla-Dave, A., Sasankan, N., and Xiao, Y.: 'Library of deep-learning image segmentation and outcomes model-implementations', Physica Medica, 2020, 73, pp. 190-196
17 Stern, U., He, R., and Yang, C.-H.: 'Analyzing animal behavior via classifying each video frame using convolutional neural networks', Scientific reports, 2015, 5, (1), pp. 1-13
18 Naito, T., Nagashima, Y., Taira, K., Uchio, N., Tsuji, S., and Shimizu, J.: 'Identification and segmentation of myelinated nerve fibers in a cross-sectional optical microscopic image using a deep learning model', Journal of neuroscience methods, 2017, 291, pp. 141-149
19 Portnoi, T., Yala, A., Schuster, T., Barzilay, R., Dontchos, B., Lamb, L., and Lehman, C.: 'Deep learning model to assess cancer risk on the basis of a breast MR image alone', American Journal of Roentgenology, 2019, 213, (1), pp. 227-233
20 Newell, A., Yang, K., and Deng, J.: 'Stacked hourglass networks for human pose estimation', in Editor (Ed.)^(Eds.): 'Book Stacked hourglass networks for human pose estimation' (Springer, 2016, edn.), pp. 483-499
21 Insafutdinov, E., Pishchulin, L., Andres, B., Andriluka, M., and Schiele, B.: 'Deepercut: A deeper, stronger, and faster multi-person pose estimation model', in Editor (Ed.)^(Eds.): 'Book Deepercut: A deeper, stronger, and faster multi-person pose estimation model' (Springer, 2016, edn.), pp. 34-50
22 Nath, T., Mathis, A., Chen, A.C., Patel, A., Bethge, M., and Mathis, M.W.: 'Using DeepLabCut for 3D markerless pose estimation across species and behaviors', Nature protocols, 2019, 14, (7), pp. 2152-2176
23 Pun, P., Brown, J., Cobb, T., Wessells, R.J., and Kim, D.H.: 'Navigation of a Freely Walking Fruit Fly in Infinite Space Using a Transparent Omnidirectional Locomotion Compensator (TOLC)', Sensors, 2021, 21, (5), pp. 1651
24 O'Mara, M.T., Wikelski, M., and Dechmann, D.K.: '50 years of bat tracking: device attachment and future directions', Methods in Ecology and Evolution, 2014, 5, (4), pp. 311-319
25 Robinson, J.L., and Jones, I.L.: 'An experimental study measuring the effects of a tarsus-mounted tracking device on the behaviour of a small pursuit-diving seabird', Behaviour, 2014, 151, (12-13), pp. 1799-1826
26 Lee, S., Waugh, B., O'Dell, G., Zhao, X., Yoo, W.-S., and Kim, D.H.: 'Predicting Fruit Fly Behaviour using TOLC device and DeepLabCut', in Editor (Ed.)^(Eds.): 'Book Predicting Fruit Fly Behaviour using TOLC device and DeepLabCut' (IEEE, 2021, edn.), pp. 1-6
27 Haralick, R.M., Joo, H., Lee, C.-N., Zhuang, X., Vaidya, V.G., and Kim, M.B.: 'Pose estimation from corresponding point data', IEEE Transactions on Systems, Man, and Cybernetics, 1989, 19, (6), pp. 1426-1446
28 He, K., Zhang, X., Ren, S., and Sun, J.: 'Deep residual learning for image recognition', in Editor (Ed.)^(Eds.): 'Book Deep residual learning for image recognition' (2016, edn.), pp. 770-778
29 Simonyan, K., and Zisserman, A.: 'Very deep convolutional networks for large-scale image recognition', arXiv preprint arXiv:1409.1556, 2014
30 Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V., and Rabinovich, A.: 'Going deeper with convolutions', in Editor (Ed.)^(Eds.): 'Book Going deeper with convolutions' (2015, edn.), pp. 1-9
31 Iandola, F.N., Han, S., Moskewicz, M.W., Ashraf, K., Dally, W.J., and Keutzer, K.: 'SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and< 0.5 MB model size', arXiv preprint arXiv:1602.07360, 2016