Improving pain assessment in mice and rats with advanced videography and computational approaches.

Nathan T. Fried[1], **Alexander Chamessian**[2,3,4], **Mark J. Zylka**[5], **Ishmail Abdus-Saboor**[6,*]

[1]Department of Biology, Rutgers University-Camden, Camden, NJ

[2]Department of Neurology, Washington University School of Medicine, St. Louis, MO

[3]Department of Anesthesiology, Washington University School of Medicine, St. Louis, MO

[4]Washington University Pain Center, Washington University School of Medicine, St. Louis, MO

[5]Department of Cell Biology and Physiology, UNC Neuroscience Center, The University of North Carolina at Chapel Hill, Chapel Hill, NC

[6]Department of Biology, University of Pennsylvania, Philadelphia, PA

1. Introduction

Accurately measuring pain in humans and rodents is essential to unravel the neurobiology of pain and discover effective pain therapeutics. However, given its inherently subjective nature, pain is nearly impossible to objectively assess. In the clinic, patients can articulate their pain experience using questionnaires and pain scales [8, 13, 20], but self-reporting can be unreliable due to various psychological and social influences or difficulties for some patients to verbalize their experience (e.g., infants, toddlers, those with neurodevelopmental disorders) [7] [25]. At the bench, these challenges are even more daunting as researchers rely on the behaviors of rodents to measure pain or pain relief. Given this, there is a growing realization among pain researchers, clinicians, and funding entities, that these traditional approaches of assessing pain in rodents may be flawed. Importantly, these flaws may have contributed to several failed drugs that initially showed promise as analgesics and point towards inconsistencies in our understanding of basic pain neurobiology [5, 17, 28].

This has prompted the field to seek new and more reliable ways to measure pain in rodents. In concert with these efforts, behavioral neuroscientists across several fields are developing new tools to improve their own behavioral assays. In this review, we describe some of these efforts and provide background for the wide adoption of these new tools by the pain research field to speed the translation of basic science findings to the clinic.

*Corresponding author: Ishmail Abdus-Saboor, Ph.D., University of Pennsylvania, Department of Biology, 3740 Hamilton Walk, 319 Leidy Laboratories, 215-898-0144, ishmail@sas.upenn.edu.

2. Sub-second analysis of acute pain assays aids in quantifying the sensory experience

The predominant measurement for pain sensation in rodents is the paw withdrawal from a noxious thermal, chemical, or mechanical stimulus. Features of the withdrawal response such as latency or frequency occur over the span of a few seconds and can be easily scored by novice experimenters. However, meaningful features that occur on a millisecond timescale go unnoticed with the unaided eye. To circumvent this, researchers have turned towards high-speed video imaging, revealing that millisecond behavioral features can be resolved and that these features contain meaningful information about the animals' internal pain state.

In one study, two of us (I.A-S, N.T.F) imaged the behavior of freely moving mice in small enclosures at 500–1,000 fps to identify the differences between sub-second movements induced by noxious versus innocuous stimulation of the hind paw [1]. Through statistical analysis and machine learning, we found that three features (paw height, paw velocity, and a combinatorial score of pain-related behaviors including orbital tightening, paw shaking, paw guarding, and jumping) could reliably distinguish withdrawals induced by noxious versus innocuous stimuli. This opened the possibility of creating a "mouse pain scale" which used a principle component analysis to combine the three features into a single number (similar to the clinically used zero-to-ten scale). We further demonstrated the utility of this pain scale by determining the sensation evoked by three von Frey hair (VFH) filaments (0.6 g, 1.4 g, and 4.0 g) and found that only the 4.0 g induced a pain-like withdrawal, while the 0.6 g and 1.4 g filaments induced a withdrawal more similar to that seen with innocuous stimuli. Finally, we examined where along this pain scale optogenetic activation of two different nociceptor populations, *Trpv1*^{Cre} and *Mrgprd*^{Cre}, would fall. Predictably, activation of *Trpv1*-lineage neurons induced a withdrawal within the pain domain. However, activation of the non-peptidergic *Mrgprd* nociceptors led to a withdrawal within the non-pain domain. Although historically considered a nociceptor population, data from other reports also suggest these neurons are not sufficient to transmit pain signals at baseline [4; 10; 15; 26].

Other groups have also extracted meaningful information from sub-second behaviors. One group identified that the latency of response to noxious stimuli was shorter (50–180 ms) than to innocuous stimuli (220–320 ms) [6]. Another group resolved sub-second movement of whisker vibrissae, body, tail, and hind paw movement to identify that an animal's posture can impact the latency of noxious-related movements, suggesting circuit-level inhibitory control of the behavior [9]. One study demonstrated that even using a lower frame rate (240 fps) with an Apple iPhone6 camera could resolve sub-second paw withdrawal, paw guarding, and jumping [3]. Collectively, these studies reveal that meaningful data can be extracted from enhanced temporal resolution of stimulus-evoked pain assays, allowing for a better approximation of an animal's pain state.

3. Considerations for using high-speed videography to capture sub-second pain behaviors

Moving from webcams, camcorders, or smartphones to dedicated high-speed cameras to study pain requires some additional considerations. Although financial considerations rank high on the list, the costs for cameras that can record around 1k fps are dropping. Another consideration is the lighting used to capture the videos. Most newer cameras can be activated with infrared lighting that will not overtly perturb the animal's behavior. Newer high-speed cameras also produce minimal sound, as to not disturb the animal. However, we still recommend habituating animals to any sounds a camera generates. Lastly, an important consideration is the hard drive space available on the camera itself, since high-speed videos can be several GB and most high-speed cameras must first save a video on internal space before transferring to external hard drives. We have found that at least 4–8 GB of internal memory on a high-speed monochrome camera is necessary for 5–10 seconds of high resolution recording (1024×1024 pixels) of behavior – which allows sufficient temporal resolution to capture sub-second and multi-second pain behaviors. With these improvements and reductions in costs, high-speed imaging is accessible enough that more labs in the field should consider incorporating high-speed imaging when evaluating paw withdrawal as an endpoint since such recordings can be used to better define the animal's behavioral state when the paw withdrawal occurs.

4. Supervised automated tracking and pose estimation with machine learning and deep neural networks to increase pain assessment workflow

High-speed imaging with manual tracking allows researchers to record behavior at millisecond resolution timescales, but the process of manually scoring and tracking these movements of multiple body areas across potentially thousands of frames or hours of data can be laborious and influenced by experimenter bias or methodological inconsistencies. Several research groups are developing new ways of automatically tracking animal posture (pose estimation) and the movements of body regions of interest (bROI) in animals ranging from invertebrates to vertebrates [11].

One new platform, DeepLabCut, has already been widely adopted across the behavioral neuroscience community [22] (Table 1). This open source system uses deep neural networks to perform pose estimation without the need of placing reflective markers on the animals. DeepLabCut was built from an earlier pose estimation algorithm called DeeperCut, which used thousands of labeled datasets to accurately track body parts of humans engaged in diverse tasks from rowing to throwing a baseball [16]. The goal of DeepLabCut was to modify the existing DeeperCut algorithm to achieve human-level tracking accuracy in animals while also reducing the amount of computer training. The use case of this software relies on an experimenter manually identifying all bROI in as few as 200 frames. The software then uses those frames to train itself to identify the same bROI throughout the entire video or other videos with similar camera angles that contain one or more animals. Thus, after the initial 200 frames of training, the software can track animal movement in all future experiments.

Another deep-learning pose estimation program, LEAP [23], is also gaining traction in the behavioral neuroscience community (Table 1). Initially developed for Drosophila, LEAP moves beyond DeepLabCut by introducing a graphical user interphase (GUI) to label body features. Similar to DeepLabCut, LEAP uses a convolutional neural network capable of making predictions of bROI in a given image. With LEAP, users only need to label the first 10 images of a dataset for training, which can be done in one hour. The authors demonstrate that LEAP can faithfully track bROI in both flies or freely behaving mice. Similar to DeepLabCut, LEAP is also free and open source with online instructions. At the time of this publication, a new version of LEAP, SLEAP (Social LEAP Estimates Animal Poses), has been announced but not yet published and will include the ability to track multiple animals in a single video.

Importantly, automated tracking of behavioral features can also be applied to the detection of spontaneous pain in rodents. The Mogil lab previously developed a mouse grimace scale (MGS) based on changes in facial expressions induced by noxious stimuli of moderate duration (10 min – 4 h) [18]. However, the MGS relies on humans scoring a small number of frames in videos, making the MGS labor intensive and subject to variability across research groups. To overcome this, the Mogil and Zylka labs built a new platform for automated scoring of the MGS (aMGS), using a convolutional neural network [24] (Table 1). The aMGS was trained with nearly 6,000 facial images of albino mice either in pain or not in pain for testing and validation. Using a post-operative pain assay, the aMGS was able to accurately predict pain, and relief of pain, with an accuracy equal to that of a human scorer. As automation of the MGS platform continues to develop, this technology should become a mainstay platform used in preclinical pain and analgesic assessment.

5.   Unsupervised identification and measurement of novel pain-related behavioral features

DeepLabCut and LEAP markerless tracking software allow a researcher to easily track bROI and then perform analysis on the subsequent data. However, these software platforms rely on the user identifying movement features of bROI. In some cases, there may be rich information not easily recognizable to humans. Thus, some groups are working to develop deep learning technologies that can automatically identify important features an experimenter should focus on. In these instances, instead of a user identifying a movement feature, deep learning software can identify previously unrecognized movement features that contain important information to distinguish between experimental groups.

The Datta lab has applied this approach to the development of a platform, MoSeq, to decode spontaneous mouse behaviors within an open field using three-dimensional (3D) depth imaging and three separate cameras (Table 1). Using this, they revealed that normal mouse behavior is structured into 60 different movement blocks (e.g., darting, freezing, low/high rearing, etc) of roughly 350 ms in length with mice frequently transitioning between these blocks [27]. They coined these movement blocks "behavioral syllables", similar to the syllables used to create words in language. Importantly, these syllables were not identified or defined by the experimenter. Instead, they were classified automatically by deep neural

networks. As a proof of principle, researchers in the Datta lab re-evaluated the walking phenotype of mouse mutants for a gene important for fluid locomotion. Although previous kinematic assessment identified abnormal motor movement patterns, this new platform uncovered a so-called 'waddle' walking phenotype in both homozygous and heterozygous mutant mice – a phenotype missed by other scoring techniques[2; 12; 19; 21].

While powerful, MoSeq is not a turn-key solution and requires substantial code customization and a fulltime computer scientist. To make unsupervised behavioral analysis more accessible, the new software suite, B-SOID, developed by the Yttri lab, was created in which the automated tracking capabilities of DeepLabCut were combined with feature extraction similar to MoSeq [14] (Table 1). The current iteration of B-SOID tracks six body parts and categorizes those positions into seven behaviors that have been manually defined in the software: pause, rear, groom, sniff, locomote, orient left, orient right. Although fairly general, these behaviors enable researchers to easily and automatically identify trends and sequences in the behaviors associated with whatever stimulus or condition is applied. Alternatively, B-SOID also offers the ability to segment statistically different behaviors based entirely on the data without any manual definition from the user. This feature of the software is perhaps most promising, as it will enable pain researchers to discover novel, non-obvious behaviors that can provide insight into the rich experience of pain. One current limitation is that B-SOID only accepts video recorded from beneath the animal. Many behaviors that pain researchers care about, like itching or scratching, are best captured either from above, the side, or a combination of multiple views demonstrating the need for B-SOID to expand to include multi-angle analysis. Nevertheless, B-SOID represents a significant advance in terms of ease of use and should enable pain researchers with minimal programming skills to take advantage of unsupervised behavioral analysis.

6. Future Directions

It should be noted that machine learning-based tools for pain research are still in their infancy. We need a greater understanding of accuracy, how well different tools predict analgesic efficacy across a broad range of drugs and pain models, concordance between tools and labs, and ease of adoption for pain researchers. Major barriers still exist despite open source sharing of new analytic pipelines. Most pain research labs lack the personnel to install, troubleshoot, and operate these somewhat finicky software packages. Efforts must be expanded to simplify the user interface and operability to facilitate broader use. Still, the combination of markerless tracking for automated pose estimation and deep learning networks for feature extraction will undoubtedly help the field identify new pain-related behaviors at both sub-second and multi-minute time scales while making already established behaviors easier and more reliable to measure. It is possible that researchers may identify nuances to rodent pain-behavior by tracking an eye, ear, paw, or even a single digit, vastly improving upon our ability to use rodents in pain research.

While the technologies detailed here will undoubtedly increase an experimenter's accuracy at measuring acute pain, it remains uncertain how well each of these technologies will perform in measuring chronic states, such as neuropathic pain. Since animals in chronic pain states often guard their injured limbs or show reductions in mounting robust motor outputs,

this may be a rate-limiting factor since some of the platforms described above rely on animals performing quick and intense movements (Figure 1). Therefore, it is possible that an advanced version of aMGS, and the unsupervised learning platforms with the capacity to detect spontaneous pain behaviors such as MoSeq or B-SOID, may be the most reliable in detecting subtle phenotypes that define the pain states during neuropathic or inflammatory pain.

To end where we began, a very small number of basic science findings in preclinical pain models are translated into the clinic as novel pain therapeutics. Therefore, another future direction could be for labs to perform pain behavior testing in rodents with some of the technologies described here, and test both known analgesics and drugs that showed painkilling promise in rodents but failed subsequently in clinical trials [29]. What if increased resolution in pain measurement in rodents could save us valuable time and energy by not pursuing the wrong targets? Only time will tell the true value in using these newer tools to study pain in mice and rats, but if our predictions are correct, the field has reason to be optimistic.

Acknowledgments

References:

[1]. Abdus-Saboor I, Fried NT, Lay M, Burdge J, Swanson K, Fischer R, Jones J, Dong P, Cai W, Guo X, Tao YX, Bethea J, Ma M, Dong X, Ding L, Luo W. Development of a Mouse Pain Scale Using Sub-second Behavioral Mapping and Statistical Modeling. Cell Rep 2019;28(6):1623–1634 e1624. [PubMed: 31390574]

[2]. Andre E, Conquet F, Steinmayr M, Stratton SC, Porciatti V, Becker-Andre M. Disruption of retinoid-related orphan receptor beta changes circadian behavior, causes retinal degeneration and leads to vacillans phenotype in mice. EMBO J 1998;17(14):3867–3877. [PubMed: 9670004]

[3]. Arcourt A, Gorham L, Dhandapani R, Prato V, Taberner FJ, Wende H, Gangadharan V, Birchmeier C, Heppenstall PA, Lechner SG. Touch Receptor-Derived Sensory Information Alleviates Acute Pain Signaling and Fine-Tunes Nociceptive Reflex Coordination. Neuron 2017;93(1):179–193. [PubMed: 27989460]

[4]. Beaudry H, Daou I, Ase AR, Ribeiro-da-Silva A, Seguela P. Distinct behavioral responses evoked by selective optogenetic stimulation of the major TRPV1+ and MrgD+ subsets of C-fibers. Pain 2017;158(12):2329–2339. [PubMed: 28708765]

[5]. Berge OG. Predictive validity of behavioural animal models for chronic pain. Br J Pharmacol 2011;164(4):1195–1206. [PubMed: 21371010]

[6]. Blivis D, Haspel G, Mannes PZ, O'Donovan MJ, Iadarola MJ. Identification of a novel spinal nociceptive-motor gate control for Adelta pain stimuli in rats. Elife 2017:e23584. [PubMed: 28537555]

[7]. Booker S, Herr K. The State-of-"Cultural Validity" of Self-Report Pain Assessment Tools in Diverse Older Adults. Pain Med 2015;16(2):232–239. [PubMed: 25219949]

[8]. Boonstra AM, Schiphorst Preuper HR, Balk GA, Stewart RE. Cut-off points for mild, moderate, and severe pain on the visual analogue scale for pain in patients with chronic musculoskeletal pain. Pain 2014;155(12):2545–2550. [PubMed: 25239073]

[9]. Browne LE, Latremoliere A, Lehnert BP, Grantham A, Ward C, Alexandre C, Costigan M, Michoud F, Roberson DP, Ginty DD, Woolf CJ. Time-Resolved Fast Mammalian Behavior Reveals the Complexity of Protective Pain Responses. Cell Rep 2017;20(1):89–98. [PubMed: 28683326]

[10]. Cavanaugh DJ, Lee H, Lo L, Shields SD, Zylka MJ, Basbaum AI, Anderson DJ. Distinct subsets of unmyelinated primary sensory fibers mediate behavioral responses to noxious thermal and mechanical stimuli. Proceedings of the National Academy of Sciences 2009;106(22):9075–9080.

[11]. Datta SR, Anderson DJ, Branson K, Perona P, Leifer A. Computational Neuroethology: A Call to Action. Neuron 2019;104(1):11–24. [PubMed: 31600508]

[12]. Eppig JT, Blake JA, Bult CJ, Kadin JA, Richardson JE, Grp MGD. The Mouse Genome Database (MGD): facilitating mouse as a model for human biology and disease. Nucleic Acids Research 2015;43(D1):D726–D736. [PubMed: 25348401]

[13]. Ferreira-Valente MA, Pais-Ribeiro JL, Jensen MP. Validity of four pain intensity rating scales. Pain 2011;152(10):2399–2404. [PubMed: 21856077]

[14]. Hsu A, Yttri E. B-SOiD: An Open Source Unsupervised Algorithm for Discovery of Spontaneous Behaviors. bioRxiv 2019.

[15]. Huang T, Lin SH, Malewicz NM, Zhang Y, Zhang Y, Goulding M, LaMotte RH, Ma Q. Identifying the pathways required for coping behaviours associated with sustained pain. Nature 2019;565(7737):86–90. [PubMed: 30532001]

[16]. Insafutdinov E, Pishchulin L, Andres B, Andriluka M, Schiele1 B. DeeperCut: A Deeper, Stronger, and Faster Multi-Person Pose Estimation Model arXiv 2016.

[17]. Kissin I. The development of new analgesics over the past 50 years: a lack of real breakthrough drugs. Anesth Analg 2010;110(3):780–789. [PubMed: 20185657]

[18]. Langford DJ, Bailey AL, Chanda ML, Clarke SE, Drummond TE, Echols S, Glick S, Ingrao J, Klassen-Ross T, Lacroix-Fralish ML, Matsumiya L, Sorge RE, Sotocinal SG, Tabaka JM, Wong D, van den Maagdenberg AM, Ferrari MD, Craig KD, Mogil JS. Coding of facial expressions of pain in the laboratory mouse. Nat Methods 2010;7(6):447–449. [PubMed: 20453868]

[19]. Liu H, Kim SY, Fu Y, Wu X, Ng L, Swaroop A, Forrest D. An isoform of retinoid-related orphan receptor beta directs differentiation of retinal amacrine and horizontal interneurons. Nat Commun 2013;4:1813. [PubMed: 23652001]

[20]. Main CJ. Pain assessment in context: a state of the science review of the McGill pain questionnaire 40 years on. Pain 2016;157(7):1387–1399. [PubMed: 26713423]

[21]. Masana MI, Sumaya IC, Becker-Andre M, Dubocovich ML. Behavioral characterization and modulation of circadian rhythms by light and melatonin in C3H/HeN mice homozygous for the RORbeta knockout. Am J Physiol Regul Integr Comp Physiol 2007;292(6):R2357–2367. [PubMed: 17303680]

[22]. Mathis A, Mamidanna P, Cury KM, Abe T, Murthy VN, Mathis MW, Bethge M. DeepLabCut: markerless pose estimation of user-defined body parts with deep learning. Nat Neurosci 2018;21(9):1281-+. [PubMed: 30127430]

[23]. Pereira TD, Aldarondo DE, Willmore L, Kislin M, Wang SSH, Murthy M, Shaevitz JW. Fast animal pose estimation using deep neural networks. Nature Methods 2019;16(1):117-+. [PubMed: 30573820]

[24]. Tuttle AH, Molinaro MJ, Jethwa JF, Sotocinal SG, Prieto JC, Styner MA, Mogil JS, Zylka MJ. A deep neural network to assess spontaneous pain from mouse facial expressions. Mol Pain 2018;14.

[25]. von Baeyer CL. Children's self-reports of pain intensity: scale selection, limitations and interpretation. Pain Res Manag 2006;11(3):157–162. [PubMed: 16960632]

[26]. Vrontou S, Wong AM, Rau KK, Koerber HR, Anderson DJ. Genetic identification of C fibres that detect massage-like stroking of hairy skin in vivo. Nature 2013;493(7434):669-+. [PubMed: 23364746]

[27]. Wiltschko AB, Johnson MJ, Iurilli G, Peterson RE, Katon JM, Pashkovski SL, Abraira VE, Adams RP, Datta SR. Mapping Sub-Second Structure in Mouse Behavior. Neuron 2015;88(6):1121–1135. [PubMed: 26687221]

[28]. Woolf CJ. Overcoming obstacles to developing new analgesics. Nat Med 2010;16(11):1241–1247. [PubMed: 20948534]

[29]. Yezierski RP, Hansson P. Inflammatory and Neuropathic Pain From Bench to Bedside: What Went Wrong? Journal of Pain 2018;19(6):571–588. [PubMed: 29307749]
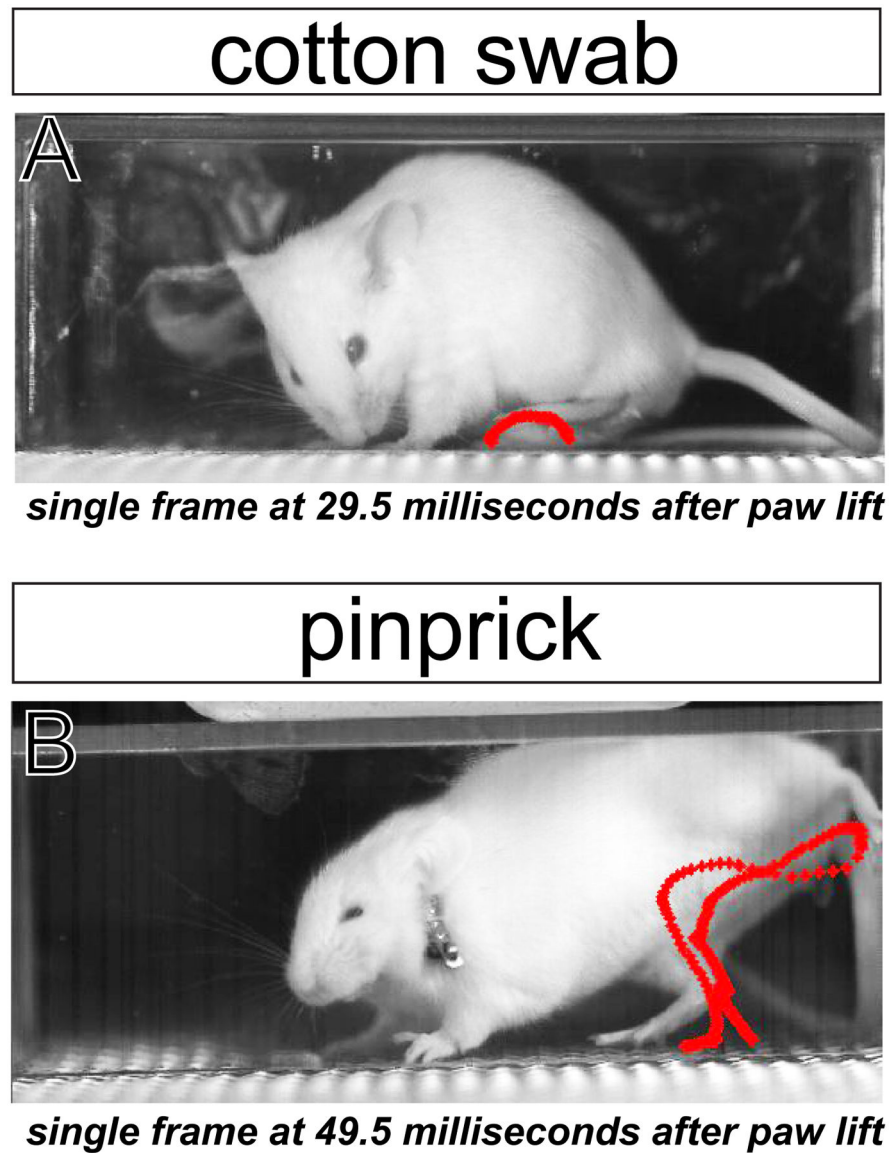
cotton swab

single frame at 29.5 milliseconds after paw lift

pinprick

single frame at 49.5 milliseconds after paw lift

Figure 1. High-speed videography reveals differences in paw movements to innocuous versus noxious stimuli.

Red paw trajectory patterns display path of stimulated paw from lift to return using automated paw tracking approaches. Orbital (eye) tightening, a MGS action unit indicative of pain, is also visible following pinprick (B), but not in response to an innocuous stimulus (A).

Table 1.

Software for markerless tracking and feature extraction of animal behavior.

Software Name	Application	Reference PMID	Open source?	Training Tutorials?	Software Link	Notes
DeepLabCut	Markerless Tracking of bROI	30127430, 31227823	Yes	Yes	https://github.com/AlexEMG/DeepLabCut	Markerless tracking removes need for reflective markers, minimal training of deep neural networks, multi-animal pose tracking integration. With lack of GUI, coding can be challenging for novice user.
LEAP	Markerless Tracking of bROI	30573820	Yes	Yes	https://github.com/talmo/leap	Markerless tracking removes need for reflective markers, minimal training of deep neural networks, GUI available. However, not able to track multiple animals in single video.
SLEAP	Markerless Tracking of bROI	Unpublished	na	na	na	Markerless tracking removes need for reflective markers, minimal training of deep neural networks, GUI available, multi-animal pose tracking integration announced.
aMGS	Automated Measurement of Mouse Grimace Scale	29546805	Yes	Yes	https://github.com/BenjaminCorvera/MGS-pipeline-distribution	Automated tracking of MGS increases workflow and decreases variability of scoring, fast detection of spontaneous behavior. Currently only works with albino mice.
MoSeq	Feature Extraction	26687221	Yes	No	Need an MTA to access software	Unsupervised extraction of important behavioral features. However, requires advanced computing ability.
B-SOID	Pose Estimation & Feature Extraction	Preprint: https://doi.org/10.1101/770271	Yes	Yes	https://github.com/YttriLab/B-SOID	Unsupervised extraction of important behavioral features, automated markerless tracking. However, current version only uses video from below, and users can not define unique behaviors to track.