Canine Video Analysis with Computer Vision

Dr. Troy Hawkins [1] Lindsey Robbins [2] Mihira Krishnaswamy [2] Anish Tiwari [2] Nolan Greene [2]
[1]Elanco Animal Health, Greenfield, IN, USA, [2] Purdue University: The Data Mine, West Lafayette, IN, USA

ABSTRACT:

Problem: Pruritus (itch) is one of several measurable criteria in the assessment of dermatitis in dogs. Watching the video of the canine behavior and annotating live is costly as it requires annotating hours of footage by hand and is prone to human error due to fatigue as well as biases.
Objective: To develop an algorithm that annotates specific canine behaviors in a video stream.
Importance: Automated video annotations will accelerate the process of canine analysis as a computer would be able to analyze hours of video. Additionally, it will reduce the possibility of errors due to human fatigue, and biases.

Introduction:

- Canine Allergic Skin Disease (cAD) affects 10-15% of the canine population
- In order to assess the effects of certain compounds on the chronic skin lesions and skin infections caused by cAD, dermatitis was induced on the study's subjects
- Pruritus(itching behavior) was observed and annotated by human observers through security camera footage during the study
- Activity monitors were used to indicate when movement occurred in the footage
- Activity monitors are not able to specify/differentiate behaviors as human observers are

Figure 1 A dog being treated for Pruritis.

- Watching the video of canine behavior and annotating live requires looking at hours of footage
- It is prone to human error due to fatigue and other factors which in turn results in a lowered consistency in accuracy
- To improve how pruritus is evaluated we worked on developing an algorithm that annotates specific canine behaviors in a video stream and can eventually be utilized by activity monitors

What is Computer Vision?

- Computer Vision is a field which deals with automating the annotation of digital images or videos.
- By utilizing machine learning, a concept which allows computers to learn how to perform tasks based on patterns in data rather than being explicitly programmed to perform those tasks, Computer Vision trains systems to annotate certain aspects of images or videos using large databases.

How is Computer Vision used in Video Analytics?

- Computer Vision has been used in a variety of fields that utilize video information. Such as facial recognition software in security systems, self-driving cars, and medical diagnostics.

What is DeepLabCut?

- DeepLabCut is a computer vision framework that tracks moving body parts across various species.
- It has the benefit of near human level identification of behaviors, using few training samples.

How does DeepLabCut work?

- Frames from the video to be analyzed are annotated indicating points of interests and joints (lines between points), signifying body parts.
- A machine learning model known as convolutional neural networks or CNN's are trained to recognize where these points and joints appear in different frames sampled from the video.
- Points and joints are reinitialized by the CNN over and over until the accuracy (computer prediction compared to human-annotated points) from the sample frames reach a certain threshold.
- The CNN can now be applied to the entire dataset and new videos (all the frames in the video) where points, joints and annotations are portrayed accurately .

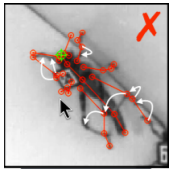


Figure 2
An example of the CNN reinitializing points and joints , (shown in white arrows)

How will DeepLabCut help with this project?

- As a canine analysis program like this has not been done before, DeepLabCut will be useful as it requires few amounts of training data, to identify canine behavior through annotated joints and points.

METHODOLOGY:

1. Data Preparation/File Conversion
- converted security camera footage (.dav format) with ffmpeg to .avi format in order to view and analyze/annotate the videos

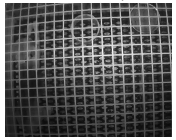
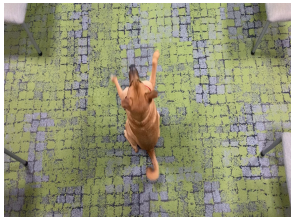
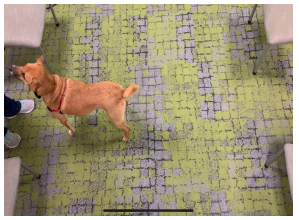


Figure 3
A screenshot of one of the videos of the beagles from Elanco.

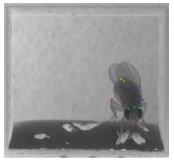
2. Video
- Took four sample videos for each behavior (spin, walking back and forth, lying down, staying still) for the purpose of analyzing different positions of the dog within the frame.
- We took the videos of Franklin (trained dog) to better understand what DeepPose can detect before we implement our algorithm on the Elanco video files.
- Each video was approximately 01:30 in duration. This length of time was chosen because it allowed us to have enough frames to train
- The 4 video files were approximately 317.5 MB in total. They were recorded in HD on iPhone XR with 1920×1080 dimensions.
- The video files were automatically compressed as they were .MOV files. When uploaded to the Scholar interface for analysis, they were converted to .AVI as needed.

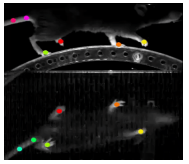


Figures 4 and 5
These are screenshotted frames from two of the different videos of Franklin. Figure 2 (left) is from the video of Franklin sitting and raising his front legs when offered a treat. Figure 3 (right) is from the video of Franklin walking back and forth when offered treats and commanded to come.



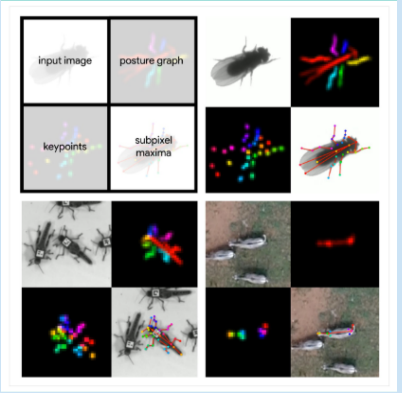
3. DeepLabCut
- Frames were extracted from the video and labeled them based on the ethogram
- 100 frames each behavior was used to train a deep neural network
- The results of the neural network, are evaluated against sample videos
- Added labels for each behavior output





CONCLUSIONS & FUTURE IMPLICATIONS:

- Our research is ongoing so we do not have conclusions to report at this time.
- Using another computer vision model to identify the points of interest instead of manually indicating them by hand is one possible optimization, suggested in the Deeppose paper.



- Developing a working algorithm that is consistently accurate could streamline the experimental process in industries which require animal behavior observation and annotation
- By adjusting the algorithm, it could be made to apply to animals other than solely canines

Acknowledgements: **Troy B. Hawkins, Ph.D.**
Distinguished Scientist and Head, Computational Sciences**, Lindsey A, Robbins**

References:

- Lindsey Robbins (PHD student), team mentor
- Toshev, et al. "DeepPose: Human Pose Estimation via Deep Neural Networks." *ArXiv.org*, 20 Aug. 2014, arxiv.org/abs/1312.4659.
- AlexEMG. "AlexEMG/DeepLabCut." *GitHub*, 27 Mar. 2020, github.com/AlexEMG/DeepLabCut
- Special thanks to Dr. Troy Hawkins and Elanco

Sponsored by Elanco Animal Health and/or one of its subsidiaries