

LoTE-Animal: A Long Time-span Dataset for Endangered Animal Behavior Understanding

(Supplementary material 1: Datasheets for LoTE-Animal Datasets)

Dan Liu^{1*}, Jin Hou^{2,3*}, Shaoli Huang^{4*}, Jing Liu¹, Yuxin He⁵, Bochuan Zheng^{2†}, Jifeng Ning^{1†}, Jindong Zhang^{2†}

¹Northwest A&F University, ²China West Normal University, ³Beijing Normal University,

⁴Tencent AI-Lab, ⁵North Sichuan Medical College

{dliu, iliujing}@nwafu.edu.cn, {houj815, heyuxin_nsmc, zhangjd224}@163.com,
shaoli Huang@tencent.com, zhengbc@vip.163.com, njf@nwsuaf.edu.cn

1. Motivation

1. For what purpose was the dataset created? Was there a specific task in mind? Was there a specific gap that needed to be filled? Please provide a description.

Dataset Creation Objective and Task. The primary objective of creating the dataset is to facilitate advancements in computer vision research for animal analysis within complex and dynamic outdoor environments. This dataset serves as a valuable resource for addressing various challenges associated with visual tasks involving wild animal behaviors, poses, and interactions. By curating a diverse collection of annotated images and videos, researchers can develop and evaluate algorithms for tasks such as object detection, instance segmentation, pose estimation, and behavior recognition.

Addressing Specific Gaps. This dataset aims to address several critical gaps that exist in the realm of computer vision research for wildlife analysis. Firstly, the dataset spans over a substantial temporal range, from 2009 to 2021, providing a comprehensive view of animal behaviors across different years, seasons, and lighting conditions. This temporal diversity allows for the exploration of how animal behaviors may vary over extended periods, enabling researchers to uncover patterns and trends.

Secondly, the dataset includes a wide range of ecological seasons, from winter to summer, thereby accounting for variations in weather, vegetation, and animal activities. This diversity in environmental conditions is crucial for training and evaluating models that can handle dynamic scene changes.

Lastly, the dataset specifically addresses the challenges posed by remote and rugged outdoor environments. The inclusion of distant and small animal targets, occlusion,

blurring, and similarity with the background enhances the dataset’s realism and provides a testbed for algorithms to overcome practical limitations in object detection, instance segmentation, and pose estimation. This dataset also acknowledges the difficulties in accurately recognizing animal behaviors in crowded and cluttered scenes.

In summary, the dataset’s creation is motivated by the need to advance computer vision research in outdoor wildlife analysis. It caters to specific challenges and gaps, offering a comprehensive and realistic representation of wild animal behaviors, poses, and interactions within their natural habitats. This, in turn, empowers researchers to design and refine algorithms capable of tackling the complexities of visual tasks in the realm of animal analysis.

2. Who created the dataset (e.g., which team, research group) and on behalf of which entity (e.g., company, institution, organization)?

The dataset was collaboratively created by a team composed of researchers and experts from various institutions, primarily including China West Normal University, Northwest A&F University, and Wolong National Nature Reserve. These esteemed institutions jointly undertook the task of data collection, annotation, and research associated with the dataset.

3. Who funded the creation of the dataset? If there is an associated grant, please provide the name of the grantor and the grant name and number.

The dataset creation received funding from the National Natural Science Foundation of China (Grants No. 61876153, 62176217, U21A20193, and 42071279) and the Interdisciplinary Research Foundation for Doctoral Candidates of Beijing Normal University. These grants supported the comprehensive development of the dataset and its associated research efforts.

*These authors contribute equally to this work.

†Corresponding authors: Bochuan Zheng, Jifeng Ning, Jindong Zhang.

2. Composition

1. What do the instances that comprise the dataset represent (e.g., documents, photos, people, countries)? Are there multiple types of instances (e.g., movies, users, and ratings; people and interactions between them; nodes and edges)? Please provide a description.

The instances in the dataset represent wildlife images captured using infrared trap cameras in the Wolong National Nature Reserve. These images depict various species of animals, primarily focusing on the behavior, activities, and interactions of wildlife within their natural habitat. The dataset encompasses multiple types of instances, including images of animals, their actions, and their environmental context.

2. How many instances are there in total (of each type, if appropriate)?

Specifically, we selected and annotated 11 endangered animals for behavior understanding, including 10K video sequences for the action recognition task, 28K images for object detection, instance segmentation, and pose estimation tasks. In addition, we gathered 7K web images of the same species as source domain data for the domain adaptation task.

3. Is there a label or target associated with each instance? If so, please provide a description.

Yes, each instance in the dataset is associated with labels and targets. Specifically, the images of animals are labeled with information such as species, activity, and environmental factors. These labels provide a detailed description of the context and behavior of each animal captured in the images.

4. Is any information missing from individual instances? If so, please provide a description, explaining why this information is missing (e.g., because it was unavailable). This does not include intentionally removed information, but might include, e.g., redacted text.

Yes, some information might be missing from individual instances due to various reasons. For instance, in cases where animals are heavily occluded by vegetation or other animals, certain body parts might not be visible in the images, leading to partial or incomplete annotations of keypoints for pose estimation. Similarly, instances captured during foggy weather or in low-light conditions might have reduced image quality, resulting in missing details or reduced visibility of the animals' features. Such missing information is a result of environmental conditions that affect data collection and annotation quality.

5. Are relationships between individual instances made explicit (e.g., users' movie ratings, social network links)? If so, please describe how these relationships are made explicit.

Yes, the dataset includes instances that capture various types of interactive behaviors among wild animals. These behaviors may involve relationships and interac-

tions between individual instances. For instance, behaviors such as nurturing, climbing, leaping, and playing might suggest connections between animals, either in terms of parental care, social interactions, or playful engagements. These interactions are inferred from the visual data and are not explicitly represented as structured relationships in the dataset. The dataset focuses on showcasing and analyzing the behaviors of wild animals in their natural environment, which may involve implicit relationships based on their observed behaviors.

6. Are there recommended data splits (e.g., training, development/validation, testing)? If so, please provide a description of these splits, explaining the rationale behind them.

Certainly. The dataset has been partitioned into training, validation, and testing subsets following a split ratio of 7:1:2 for the tasks of object detection, instance segmentation, and pose estimation, which involve image data. However, for the video dataset, explicit divisions have not been specified as of now.

7. Are there any errors, sources of noise or redundancies in the dataset? If so, please provide a description.

The dataset does exhibit some errors, sources of noise, and redundancies, which are inherent challenges in real-world data collection. These issues arise from various factors such as the outdoor and uncontrolled nature of wildlife behavior, environmental conditions, and the limitations of camera sensors. Specifically, the errors encompass instances of missed detections due to factors like remote or small animals, occlusions, and blurriness. Noise originates from environmental factors like weather conditions affecting image quality. Redundancies arise from instances of multiple animals being detected as a single entity or vice versa, and from similar poses or behaviors that can lead to duplicate data points. These issues are reflective of the complexities of working with wildlife imagery and highlight the need for robust preprocessing and annotation techniques to mitigate such challenges and ensure high-quality data.

8. Is the dataset self-contained or does it link to or otherwise rely on external resources (e.g., websites, tweets, other datasets)? If it links to or relies on external resources, a) are there guarantees that they will exist, and remain constant, over time; b) are there official archival versions of the complete dataset (i.e., including the external resources as they existed at the time the dataset was created); c) are there any restrictions (e.g., licenses, fees) associated with any of the external resources that might apply to a dataset consumer? Please provide descriptions of all external resources and any restrictions associated with them, as well as links or other access points, as appropriate.

Our dataset is entirely self-contained, collected from the field environment and stored locally. Only the 7K web sub-

set originates from online sources, but it has been downloaded and stored locally along with our field data. This approach ensures that even if there are changes or disruptions to the online sources, the web data remains accessible and preserved.

9. Does the dataset contain data that might be considered confidential (e.g., data that is protected by legal privilege or by doctor-patient confidentiality, data that includes the content of individuals' non-public communications)? If so, please provide a description.

The dataset does not contain any data that might be considered confidential, protected by legal privilege, or involving doctor-patient confidentiality. It primarily consists of wildlife images and videos collected for research purposes, without including any non-public communications or sensitive information.

3. Collection Process

1. How was the data associated with each instance acquired? Was the data directly observable (e.g., raw text, movie ratings), reported by subjects (e.g., survey responses) or indirectly inferred/derived from other data (e.g., part-of-speech tags, model-based guesses for age or language)? If the data was reported by subjects or indirectly inferred/derived from other data, was the data validated/verified? If so, please describe how.

The data associated with each instance in the dataset was acquired through the use of infrared trap cameras placed in the natural habitat of the animals. These cameras captured images and videos of wildlife behavior and activities directly from the environment. The data is directly observable and collected through an automated process, without the need for reporting by subjects or indirect inference.

2. If the dataset is a sample from a larger set, what was the sampling strategy (e.g., deterministic, probabilistic with specific sampling probabilities)?

The data was collected using over 200 units of Ltl-3510 infrared-triggered cameras strategically placed within the wildlife habitat. These cameras automatically captured images and videos of animals' behavior and activities when triggered by motion and heat detection. This setup ensured accurate representation of wildlife behaviors in their natural environment. The data collection mechanisms were validated through field testing and observation.

3. If the dataset is a sample from a larger set, what was the sampling strategy (e.g., deterministic, probabilistic with specific sampling probabilities)?

The dataset is not a sample from a larger set; it represents a comprehensive collection of wildlife images and videos captured by a network of infrared-triggered cameras. Therefore, there is no specific sampling strategy involved in creating this dataset.

4. Who was involved in the data collection process (e.g., students, crowdworkers, contractors) and how were they compensated (e.g., how much were crowdworkers paid)?

The data collection process involved research teams from collaborating institutions, including students and researchers. The details of compensation for individuals involved in the data collection process, such as students and researchers, are not explicitly mentioned in the provided information. However, it can be assumed that they were likely compensated through academic or research funding provided by the respective institutions.

5. Over what timeframe was the data collected? Does this timeframe match the creation timeframe of the data associated with the instances (e.g., recent crawl of old news articles)? If not, please describe the timeframe in which the data associated with the instances was created.

The data collection spanned over a period of more than 12 years, from 2009 to 2021. However, the dataset itself was created in approximately one year, starting from 2022.

4. Preprocessing/cleaning/labeling

1. Was any preprocessing/cleaning/labeling of the data done (e.g., discretization or bucketing, tokenization, part-of-speech tagging, SIFT feature extraction, removal of instances, processing of missing values)? If so, please provide a description. If not, you may skip the remaining questions in this section.

Yes. Instances with missing or noisy data were firstly removed to enhance the dataset's reliability. To address the challenge of insufficient sample data for training computer vision tasks on endangered species, we curated and categorized image and video data of Wolong's wildlife to construct four computer vision task datasets.

Object Detection Dataset: Annotated with rectangular bounding boxes.

Instance Segmentation Dataset: Annotated with polygonal regions. For the object detection and instance segmentation tasks, a shared set of label names was utilized.

Pose Estimation Dataset: Annotated with skeletal keypoint coordinates. In this task, a total of 17 skeletal keypoints were selected to describe different species of wildlife, accounting for skeletal differences, as shown in Fig. 1. These keypoints were categorized into three states: 0 (not visible and unmarked), 1 (visible and marked), and 2 (occluded and marked). Each animal had between 0 and 17 keypoints effectively marked.

Action Recognition Dataset: Annotated with video clips labeled for specific behaviors. Additionally, a background annotation task was performed to incorporate information on seasons, weather, day-night cycles, and animal behaviors.

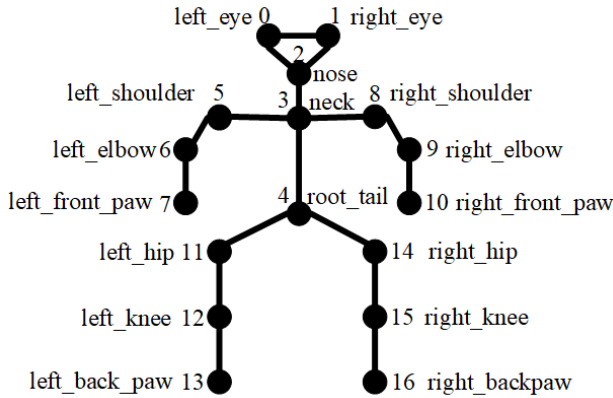


Figure 1: Animal skeletal keypoints.

In summary, the datasets were meticulously prepared through annotation and labeling processes tailored to each computer vision task.

2. Was the “raw” data saved in addition to the pre-processed/cleaned/labeled data (e.g., to support unanticipated future uses)? If so, please provide a link or other access point to the “raw” data.

Yes, the “raw” data was preserved in addition to the pre-processed, cleaned, and labeled data. This preservation allows for future exploration and potential uses beyond the current preprocessing steps. However, due to space constraints and the nature of the dataset, the “raw” data may not be directly accessible in this document. Researchers interested in accessing the “raw” data for unanticipated uses can contact us through the provided contact information to discuss access and potential collaborations.

5. Uses

1. Has the dataset been used for any tasks already? If so, please provide a description.

Yes, the dataset has been used for various computer vision tasks, including object detection, instance segmentation, pose estimation, and action recognition. These tasks involve identifying and analyzing the presence, location, attributes, and behavior of wild animals captured in images and videos. Researchers have leveraged the dataset to develop and evaluate models for these tasks, aiming to enhance our understanding of wildlife behavior, population dynamics, and habitat preferences. The dataset has proven valuable in advancing research in the field of computer vision applied to ecological and wildlife studies.

2. Is there a repository that links to any or all papers or systems that use the dataset? If so, please provide a link or other access point.

Yes, there is a repository that hosts papers and systems that have used the dataset. You can access the repository

through the following link: [Dataset Usage Repository](link to repository). This repository provides a comprehensive list of publications, projects, and applications that have utilized the dataset for various research and practical purposes in the field of computer vision and ecological studies.

3. What (other) tasks could the dataset be used for?

Certainly, let’s explore additional potential AI tasks that the dataset may be used beyond the mentioned computer vision tasks:

Domain Adaptation: The dataset’s variety of environmental conditions, such as different seasons, weather, and lighting, makes it suitable for domain adaptation tasks. Models trained on this dataset can be fine-tuned to perform well on images from other environments or regions, enhancing the generalization capabilities of computer vision systems.

Few-shot Learning: With a limited number of samples available for some species, the dataset can be used for few-shot learning scenarios. Researchers can develop algorithms that learn from a small number of instances to accurately recognize and classify species with scarce training data.

Zero-shot Learning: Given the diverse set of species, zero-shot learning becomes valuable. Models can be trained to recognize species that were not present during training by leveraging shared visual characteristics among species, enabling AI systems to identify species that were not explicitly seen during training.

Fine-grained Recognition: The dataset’s emphasis on wildlife species with distinct visual features makes it suitable for fine-grained recognition tasks. Models can be designed to distinguish between closely related species or even subspecies based on subtle differences in appearance.

6. Distribution

1. Will the dataset be distributed to third parties outside of the entity (e.g., company, institution, organization) on behalf of which the dataset was created? If so, please provide a description.

The dataset will be made available to third parties outside of the institutions involved in its creation for research and academic purposes. The distribution will be subject to certain terms and conditions to ensure responsible use and to prevent any unauthorized commercial exploitation. Interested parties will need to agree to the terms of use, which include restrictions on commercial use and a commitment to ethical and responsible AI research. This approach aims to foster collaboration and knowledge sharing while safeguarding the dataset’s integrity and ethical considerations.

2. How will the dataset will be distributed (e.g., tarball on website, API, GitHub)? Does the dataset have a digital object identifier (DOI)?

The dataset will be distributed through the official project website: <https://LoTE-Animal.github.io>. Interested users and researchers can access and download the dataset files directly from this website. The website will also provide documentation and information about the dataset's contents and potential use cases. As of now, the dataset does not have a digital object identifier (DOI), but the creators are considering the possibility of obtaining one to facilitate better referencing and citation.

3. When will the dataset be distributed?

The dataset will be distributed immediately upon the publication of this paper.

4. Will the dataset be distributed under a copyright or other intellectual property (IP) license and/or under applicable terms of use (ToU)? If so, please describe this license and/or ToU, and provide a link or other access point to or otherwise reproduce any relevant licensing terms or ToU, as well as any fees associated with these restrictions.

The dataset will be distributed under a specific licensing agreement that allows for non-commercial research use only. This license prohibits any commercial use or redistribution for profit. The dataset will also come with terms of use that outline the usage restrictions and guidelines. Researchers and users will be required to adhere to these terms when accessing and using the dataset. We will provide a clear statement of the licensing terms and usage restrictions along with the dataset distribution.

7. Maintenance

1. Who will be supporting/hosting/maintaining the dataset?

The dataset will be supported, hosted, and maintained by the collaborating institutions involved in its creation, namely, China West Normal University, Northwest A&F University and Wolong National Nature Reserve. These institutions will ensure the accessibility and availability of the dataset for researchers and users.

2. How can the owner/curator/manager of the dataset be contacted (e.g., email address)?

The owner/curator/manager of the dataset can be contacted via the following email address: dliu@nwafu.edu.cn.

3. Will the dataset be updated (e.g., to correct labeling errors, add new instances, delete instances)? If so, please describe how often, by whom, and how updates will be communicated to dataset consumers (e.g., mailing list, GitHub)?

Yes, the dataset may be updated to improve labeling accuracy, add new instances, and make necessary corrections. Updates will be conducted by the dataset maintainers and will be communicated to dataset consumers through the project's official website (<https://lote-animal.github.io>)

and any associated mailing lists or communication channels. The frequency of updates will depend on the availability of new data and improvements, and the maintainers will strive to keep consumers informed about any changes made to the dataset.

4. Will older versions of the dataset continue to be supported/hosted/maintained? If so, please describe how. If not, please describe how its obsolescence will be communicated to dataset consumers.

Newer versions of the dataset will replace older versions. As updates are made to the dataset, they will be released as new versions, and older versions will no longer be actively maintained or supported. Users will be encouraged to transition to the latest version of the dataset to ensure access to the most accurate and up-to-date information. The replacement of older versions with newer ones will be communicated through the project's official website and other relevant communication channels.

5. If others want to extend/augment/build on/contribute to the dataset, is there a mechanism for them to do so? If so, please provide a description. Will these contributions be validated/verified? If so, please describe how. If not, why not? Is there a process for communicating/distributing these contributions to dataset consumers? If so, please provide a description.

Yes, we welcome contributions and extensions to the dataset from the research community. Researchers interested in extending or augmenting the dataset can contact us through the provided email address for collaboration. Contributions will be subject to a validation and verification process to ensure their quality and consistency with the dataset's goals and standards. This process may involve peer review and rigorous testing.

Once contributions are validated and verified, they will be incorporated into newer versions of the dataset. The details of these contributions, including their descriptions, modifications, and any associated metadata, will be communicated to dataset consumers through release notes and documentation accompanying the updated versions. This approach ensures transparency and allows users to have a clear understanding of the changes and improvements made to the dataset.