Data in Brief 52 (2024) 110027

Contents lists available at ScienceDirect

Data in Brief

journal homepage: www.elsevier.com/locate/dib

Data Article

Video dataset of sheep activity for animal behavioral analysis via deep learning

Nathan A. Kelly[a], Bilal M. Khan[a,*], Muhammad Y. Ayub[b], Abir J. Hussain[c], Khalil Dajani[a], Yunfei Hou[a], Wasiq Khan[d]

[a] *School of Computer Science and Engineering, California State University San Bernardino, 5500 University Parkway, San Bernardino, CA 92407, USA*
[b] *COMSATS University Islamabad, Attock Campus, Near Officers colony, Kamra Road, Attock, Pakistan*
[c] *Department of Electrical Engineering, University of Sharjah, Sharjah, United Arab Emirates*
[d] *Faculty of Engineering and Technology, Liverpool John Moores University, Liverpool L33AF, UK*

ARTICLE INFO

Article history:
Received 24 October 2023
Revised 29 December 2023
Accepted 2 January 2024
Available online 9 January 2024

Dataset link: Video Dataset of Sheep Activity (Grazing, Running, Sitting) (Original data)
Dataset link: Video Dataset of Sheep Activity (Standing and Walking) (Original data)

Keywords:
Pattern recognition
Deep learning
Animal surveillance
Farming
Classification
Sheep activity detection

ABSTRACT

A primary dataset capturing five distinct types of sheep activities in realistic settings was constructed at various resolutions and viewing angles, targeting the expansion of the domain knowledge for non-contact virtual fencing approaches. The present dataset can be used to develop non-invasive approaches for sheep activity detection, which can be proven useful for farming activities including, but not limited to, sheep counting, virtual fencing, behavior detection for health status, and effective sheep breeding. Sheep activity classes include grazing, running, sitting, standing, and walking. The activities of individuals, as well as herds of sheep, were recorded at different resolutions and angles to provide a dataset of diverse characteristics, as summarized in Table 1. Overall, a total of 149,327 frames from 417 videos (the equivalent of 59 minutes of footage) are presented with a balanced set for each activity class, which can be utilized for robust non-invasive detection models based on computer vision techniques. Despite a decent existence of noise within the original data (e.g., segments with no sheep present, multiple sheep in single frames, multiple activities by one or more sheep in single as well as multiple frames, segments with sheep alongside other non-sheep objects), we provide

* Corresponding author.
E-mail address: Bilal.Khan@csusb.edu (B.M. Khan).

https://doi.org/10.1016/j.dib.2024.110027

original videos and the original videos' frames (with videos and frames containing humans omitted for privacy reasons). The present dataset includes diverse sheep activity characteristics and can be useful for robust detection and recognition models, as well as advanced activity detection models as a function of time for the applications.

Table 1: Summary of the dataset with hardware specifications.

Subject	Computer Science
Specific subject area	Virtual Fencing, Sheep Activity Detection, Animal Surveillance, Computer Vision, Pattern Recognition, Deep Learning
Type of data	Videos, Images
Data acquisition apparatus	Sheep with five different activities were captured via two digital camera devices with the following specifications: • **Side pose video capture:** Camera manufacturer: Apple, Phone model: iPhone XS Max, Camera: iPhone dual 12MP wide angle and telephoto cameras, Lens: 6 element lens 120° field of view, Aperture value: f/1.8, Video recording: 60fps 4K recording • **Sagittal pose video capture:** Camera manufacturer: Xiaomi, Phone model: Redmi Note 10 Pro, Camera: Quad 108MP wide angle 26mm camera, Aperture value: f/1.9, Video recording: 30fps 4K recording. Data at the scales of individual and sequential ordered frames consisted of non-sheep objects where frames that included human subjects were removed.
Data format	MOV/MP4, PNG
Description of data collection	Sheep activities of five classes (grazing, running, sitting, standing, and walking) were captured and labeled. High-resolution videos and their individual frames are provided in the dataset, which can effectively be used for developing deep learning approaches for sheep detection and behavior recognition. A total of 417 videos with 149,327 frames are presented using various angles (sagittal and side) and resolutions, conveniently organized in labeled directories for each of the five labeled classes.
Data source location	Country: Pakistan, cites: Attock, Hazro 33°55′23.5"N 72°33′07.3"E
Data accessibility	The proposed dataset is publicly available at Mendeley platform as described below: Repository name: Mendeley Data[1] Data identification number: 1. Standing and Walking classes: 10.17632/w65pvb84dg.1 2. Grazing, Sitting, and Running classes: 10.17632/h5ppwx6fn4.1 Direct URL to data: 1. Standing and Walking classes: https://data.mendeley.com/datasets/w65pvb84dg/1 2. Grazing, Sitting, and Running classes: https://data.mendeley.com/datasets/h5ppwx6fn4/1

[1] The size of the folder with 149,327 frames of high resolution is significantly above the upload limit provided by Mendeley Data Repository. The available frames along with labeled masks of sheep are retained and can be provided upon request to the corresponding authors.

1. Value of the Data

- The first-of-its-kind dataset can be useful for real-time sheep activity identification, and can be utilized for various applications such as sheep herd management, animal counting, behavioral analysis, virtual fencing, livestock farming, and animal health status updates.
- Automated sheep activity identification approaches can be developed using the proposed dataset, which can aid reduced manual labor efforts (e.g., drenching, shearing) to safer animal farming with non-invasive techniques.
- Although various approaches have been developed for sheep activity detection based on various datasets and primary videos [1–7], the proposed primary dataset with videos of high resolution can expand the available domain of knowledge, allowing for existing pretrained detection models to be finetuned as well as new approaches to be tested.
- With the availability of the code to generate labeled frames and masks from the videos in the proposed dataset, advanced approaches can be considered where timeseries forecasting models can be utilized to describe the dynamic behavior of sheep as well as high resolution multiclass sheep segmentation since the dataset is diverse and represents multiclass behavior.
- Deep learning models can be developed using the proposed dataset and made available as pretrained models for other animals for accelerated application.
- The dataset is available for public download and allows for the application of new computer vision approaches for sheep detection.

2. Objective

The proposed dataset is constructed for use in computer science, particularly the areas of pattern recognition, deep learning, and computer vision to develop sheep activity detection models that could prove to be useful non-invasive and efficient approach for livestock management, virtual fencing, health status checks, breed management, and other farming related events. Novel approaches developed via the proposed dataset can provide an alternative to continuous labor-intensive monitoring activities, leading to reduced need for constant surveillance and the labor costs associated with that practice. Furthermore, safety can be improved for the detection of unusual/unhealthy behavior of animals, allowing for timely mitigation strategies.

3. Data Description

The proposed dataset consists of two folders: one containing the original videos, and one containing the individual video frames, with those containing humans omitted for privacy reasons. A total of 417 videos are provided, averaging 8.5118 s each, resulting in a total of 146,964 frames with additional 2363 frames of miscellaneous noise or data that can be used for testing (with 41,683 total frames omitted). The videos and frames are high-resolution, with the dimensions of either 1920×1080 or 3840×2160 pixels (Table 2). Within the dataset, the subset of extracted individual frames within each class also consist of multiple sheep with >1 activities which can be used to assess the robustness of the newly developed detection techniques. Here we note that the deep learning models can be developed as multiclass detection models where each segmented sheep can be assigned a particular class given the learned model features. A randomly selected set of ten images (2 for each activity class) are shown in Fig. 1.

In the present study, the dataset was also processed to extract sheep features in the form of regions (image masks) representing the areas within each image corresponding to the taxonomy of the five classes of activities. The regions with labeled binary masks (i.e., pixels belonging to a sheep labeled with 1 and 0 otherwise) were prepared and saved as individual mask images to support the development of advanced segmentation models. Binary masks were generated in

Table 2
The framerates and resolutions of the videos for each sheep activity within the dataset.

Class	Videos at 30fps	Videos at 60fps	Still images	Videos at 3820 × 2160	Videos at 1920 × 1080	Still images at 4000 × 3000	Total Videos
Extra Activities	6	1	1	6	1	1	8
Grazing	54	36	0	88	2	0	90
Running	43	20	0	63	0	0	63
Sitting	52	47	0	96	3	0	99
Standing	61	49	0	108	2	0	110
Walking	33	14	0	47	0	0	47

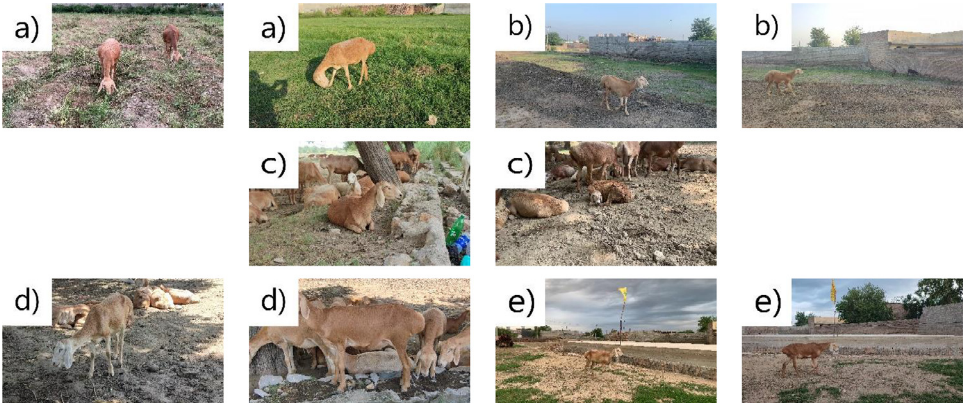


Fig. 1. Randomly selected frames representing five classes of sheep activity: **(a)** grazing, **(b)** running, **(c)** sitting, **(d)** standing, and **(e)** walking.

both transparent background (PNGs) and monochromatic bitmap formats which can be provided upon request to the corresponding authors. We also note that the frames without the existence of sheep were removed from the mask dataset, with a remaining total of 185,719 full-size frames and 1,301,119 cutouts of individual sheep. Dataset with labeled mask images can be readily used to develop advanced classification/segmentation techniques which should handle various existing challenges that mainly include: (i) lighting effects, (ii) foreshortening (distortions caused by the angle at which the sheep were recorded), (iii) aspect ratio affected by the viewing angle, (iv) occlusion due to multiple sheep present in the same frame, and (iv) object deformation (e.g., shape of the sheep varies due to different poses while sitting, walking and running).

4. Experimental Design, Materials and Methods

Acquisition of the videos for each sheep activity class followed the workflow as illustrated in Fig. 2. In the proposed approach, camera positioning at predefined locations was performed for each video session for stable and high-quality 4K video acquisition. Xiaomi's Redmi Note 10 Pro and iPhone XS Max (Table 1) were used for sagittal (4K video at 3840 × 2160 at 30fps), and side pose (1920 × 1080 at 60fps) video recording, respectively. A randomly selected sheep (or a herd of sheep) was selected in the video from two angles (i.e., sagittal and side) to record the video for an average of 8.5118 s (Table 2). Each sheep was recorded in 4K video resolution from both sagittal and side angles to include sufficient data granularity, which resulted in a total of 417

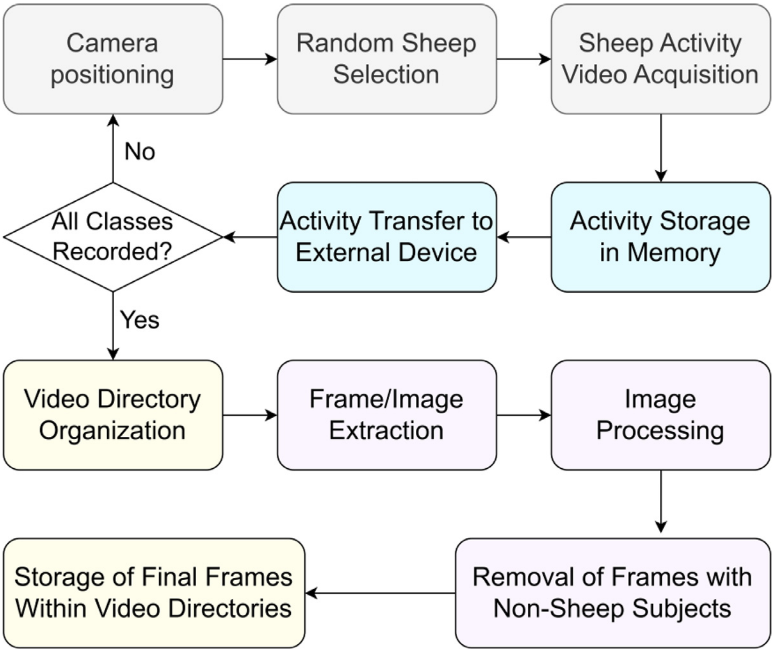
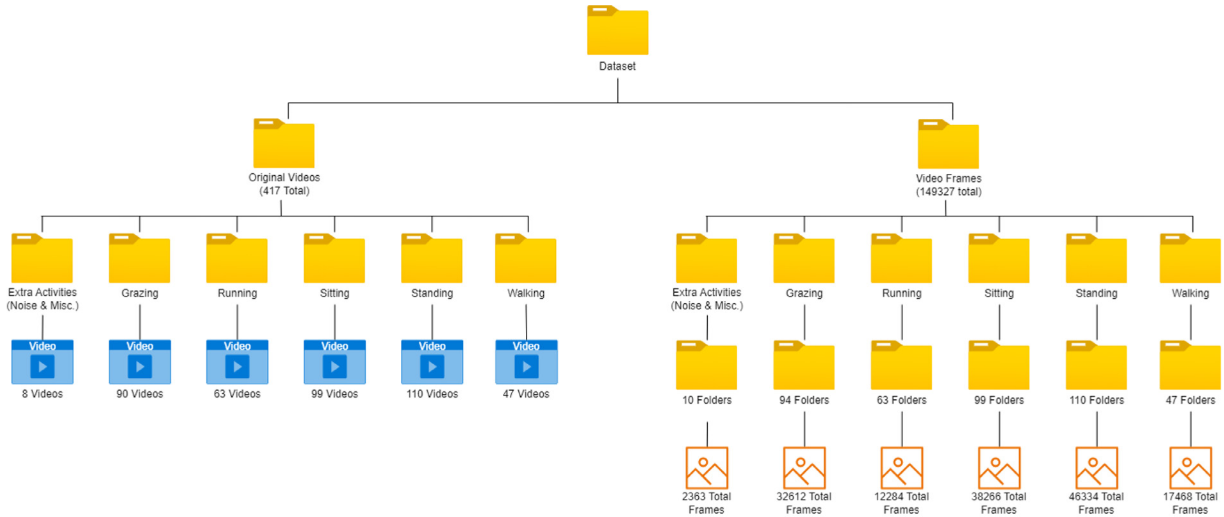


Fig. 2. Workflow for the acquisition, storage, processing, and organization of the sheep activity for each of the five classes.

videos. The activities of each sheep were stored in local memory which was then transferred to an external device. The above process continued until sufficient videos of the sheep (>70) for each class at the above two angles were acquired. Organization of these videos in the directories with correct labels was performed prior to the extraction of each frame from each video. An automated script was prepared to iteratively select each video segment and extract frames/images to store in a separate sub-directory structure (within the root directory). Here we note that while the predictive model's accuracy can be dependent on the image quality, particularly for segmentation tasks, the quality of the included images can be reduced to provide readily available data for time-sensitive modeling approaches (e.g., development and deployment of models with low computational capacity onto handheld devices for real-time tracking and segmentation). Additionally, individual videos and images were analyzed to remove those which recorded non-sheep subjects (humans) to adhere to privacy and confidentiality concerns. Consent was also acquired from the owners of the sheep herd prior to obtaining the dataset for each sheep activity. Finally, the processed images at two different resolutions were organized in a separate sub-directory at the same directory level as of the original. Final directory structure including video segments, images at two different resolutions is illustrated in Fig. 3.

Limitations

A comprehensive video set of sheep activities were captured via ground-based cameras and prepared in this article, which can be used for developing new deep learning models and enabling accelerated application of pretrained models for other similar domains. However, we note that, since herd management via Unmanned Aerial Vehicles (UAV) (including drone technologies [1,8] and networks using drone captures to detect and count animals [9]) have recently gained



(a) Video set **(b) Image set with frames extracted from videos**

Fig. 3. An illustration of the organization of the dataset's data. A total of 417 original videos separated by class are provided in the dataset (a); Individual frames from these videos were also extracted and organized per activity class, resulting in a total of 149,327 frames (b). We note here that the set of extracted frames were separated from the dataset due to the limited upload permission by the open-source Mendeley Data Repository. The extracted frames in the above format along with the labeled mask images can therefore be made available upon request to the corresponding authors.

interest, the accuracy of activity detection techniques using this proposed dataset may vary compared to the datasets obtained from UAV videos, due to the differences between the learned features of the deep learning models. The above technical barrier may, however, be overcome by using Transfer Learning approaches (transferring model parameters and utilizing minimal or small datasets) to obtain higher-accuracy detections.

We also note that the extracted individual frames and their corresponding masks for each video in the dataset are not included due to the maximum data upload limit, which required up to 2TB of storage space. We however include the code (generate_masks.py file) for generating frames and masks on a machine with the following required libraries:

- NumPy v1.24.3,
- OpenCV (cv2) v4.7.0.72,
- PyTorch (torch) v2.0.1,
- Segment Anything v1.0 [10]
- Ultralytics. v8.0.109

Videos in the proposed dataset were acquired via ground-based devices without the use of a stationary stand which allowed to capture and follow the path of the selected sheep (or herd of sheep). Moderate oscillations exist in the videos due to the above, which may also be challenging for utilizing the approaches that are based on background subtraction techniques. We however note that the primary focus of this study was to generate videos for advanced computer vision techniques that rely on extracted features (such as deep convolutional neural networks and Long Short-Term Memory (LSTM)) and are not limited detection in videos with stationary background.

Finally, the quality of image masks with respect to accurate pixelwise representation is reliant on the adopted modeling approaches of You Only Look Once (YOLO) and Segment Anything [10]. For example, 5.05% of sheep detected by YOLO on average were detected multiple times under different classifications, resulting in duplicate segmentations. Similarly, four-legged non-sheep animals can also be detected by YOLO affecting an average of only 1.4% of the detections, whereas Segment Anything showed to combine multiple overlapping sheep into a contiguous segmentation with an average misclassification rate of 4.59%. This constraint may also be apparent for other available mask generation techniques. Although manual processing of each of the 146,964 frames for mask generation may significantly hinder the process of segmentation model development, other candidate mask generation approaches can be applied depending on their accuracy.

Ethics Statement

The proposed data does not involve any human subjects, animal experiments, or data collected from social media platforms. The study for data collected from the designated devices was approved by the institution (COMSATS University Islamabad, Attock Campus, Pakistan) with the details as follows:

Approval No: FYP-EthCertificate/01-2022/CUIATK

Project Title: Virtual Fencing and Sheep Activity Detection using Machine Learning and Computer Vision

Approval Date: 28 July 2022

CRediT Author Statement

Bilal Khan: Conceptualization, Methodology, Supervision, Writing-Original draft preparation, Reviewing and Editing. **Muhammad Y. Ayub:** Video acquisition apparatus configuration, video data acquisition and transmission to external storage media and a secure cloud storage facility, Writing-Original draft preparation. **Nathan Kelly:** Data processing and preparation, Writing-Original draft preparation, visualization. **Wasiq Khan:** Methodology, Investigation, Writing-Original draft preparation, Reviewing and Editing. **Yunfei Hou:** Writing, Reviewing and Editing. **Khalil Dajani:** Supervision, Writing- Reviewing and Editing. **Abir Hussain:** Supervision, Writing- Reviewing and Editing.

Data Availability

Video Dataset of Sheep Activity (Grazing, Running, Sitting) (Original data) (Mendeley Data)
Video Dataset of Sheep Activity (Standing and Walking) (Original data) (Mendeley Data)

Acknowledgments

This material is based upon work supported by the U.S. National Science Foundation under Grant No. DUE-2322436.

Declaration of Competing Interest

None.

Supplementary Materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.dib.2024.110027.

References

[1] H. Zhao, R. Mao, M. Li, B. Li, M. Wang, SheepInst: a high-performance instance segmentation of sheep images based on deep learning, Animals 13 (8) (2023), doi:10.3390/ani13081338.
[2] F. Sarwar, A. Griffin, P. Periasamy, K. Portas, J. Law, Detecting and counting sheep with a convolutional neural network, in: 2018 15th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS), 2018, pp. 1–6, doi:10.1109/AVSS.2018.8639306.
[3] J. Kaler, J. Mitsch, J.A. Vázquez-Diosdado, N. Bollard, T. Dottorini, K.A. Ellis, Automated detection of lameness in sheep using machine learning approaches: novel insights into behavioural differences among lame and non-lame sheep, R. Soc. Open Sci. 7 (1) (2020) 190824, doi:10.1098/rsos.190824.
[4] Y. Pang, W. Yu, Y. Zhang, C. Xuan, P. Wu, Sheep face recognition and classification based on an improved MobilenetV2 neural network, Int. J. Adv. Robot. Syst. 20 (1) (2023) 17298806231152968, doi:10.1177/17298806231152969.
[5] C. Ma, X. Sun, C. Yao, M. Tian, L. Li, Research on sheep recognition algorithm based on deep learning in animal husbandry, J. Phys. Conf. Ser. 1651 (1) (2020) 12129, doi:10.1088/1742-6596/1651/1/012129.
[6] Z. Wan, F. Tian, C. Zhang, Sheep face recognition model based on deep learning and bilinear feature fusion, Animals 13 (12) (2023), doi:10.3390/ani13121957.
[7] Y. Halimu, Z. Chao, J. Sun, and X. Zhang, "Sheep herd recognition and classification based on deep learning BT - advances in natural computation, fuzzy systems and knowledge discovery," 2023, pp. 122–129.
[8] M.A. Alanezi, M.S. Shahriar, M.B. Hasan, S. Ahmed, Y.A. Sha'aban, H.R.E.H. Bouchekara, Livestock management with unmanned aerial vehicles: a review, IEEE Access 10 (2022) 45001–45028, doi:10.1109/ACCESS.2022.3168295.
[9] F. de Lima Weber, et al., Counting cattle in UAV images using convolutional neural network, Remote Sens. Appl. Soc. Environ. 29 (2023) 100900, doi:10.1016/j.rsase.2022.100900.
[10] A. Kirillov et al., "Segment Anything," *ArXiv*, vol. abs/2304.02643, 2023, [Online]. Available: https://api.semanticscholar.org/CorpusID:257952310.