

# Dairy Calf Action Recognition and Analysis Based on Spatio-Temporal Convolutional Neural Network

Po-Lin Chen, Ta-Te Lin \*

Department of Biomechatronics Engineering, National Taiwan University



## Introduction

Dairy calf behavior, which reflects their health and welfare, has traditionally been assessed through visual observation. Leveraging imaging and deep learning, our study has established a system using a spatio-temporal Convolutional Neural Network based on consecutive frames. Our study aims to enhance livestock management strategies, thereby ensuring the optimal well-being and health of dairy calves.

## Materials and Methods

### data collection

- A Wireless Imaging Device with a top-down perspective was used to capture dairy calf images.
- For action recognition network training, a sequence of 32 consecutive frames at 10 FPS is utilized.



Figure 1. Wireless Imaging Device.

### model training

- In this study, we used R(2+1)D, a Convolutional Neural Network based on ResNet. The network combines 2D and 1D convolutional operations to effectively capture spatio-temporal relationships (Tran et al., 2018).
- With this deep learning methodology, we trained on video datasets consisting of 32 frames, each with a resolution of 224 x 224. Subsequently, we categorized dairy calf behavior into six distinct classes: *Non-active Lying*, *Active Lying*, *Non-active Standing*, *Active Standing*, *Feeding*, and *Drinking*.

### data sampling and inference

- Data was periodically recorded in segments of 3.2 seconds at 1-minute intervals. The collected data was then processed with a trained spatio-temporal Convolutional Neural Network.
- The daily behavioral distribution was mapped, and further analysis was performed to quantify the total duration, number of bouts, and duration of each bout for every class.

## Results

### action recognition

- About 2,000 video segments were used for training and validation, with additional 450 reserved for the independent testing dataset.
- After training, the model achieved a top-1 accuracy of 0.94 and an F1-score of 0.93.

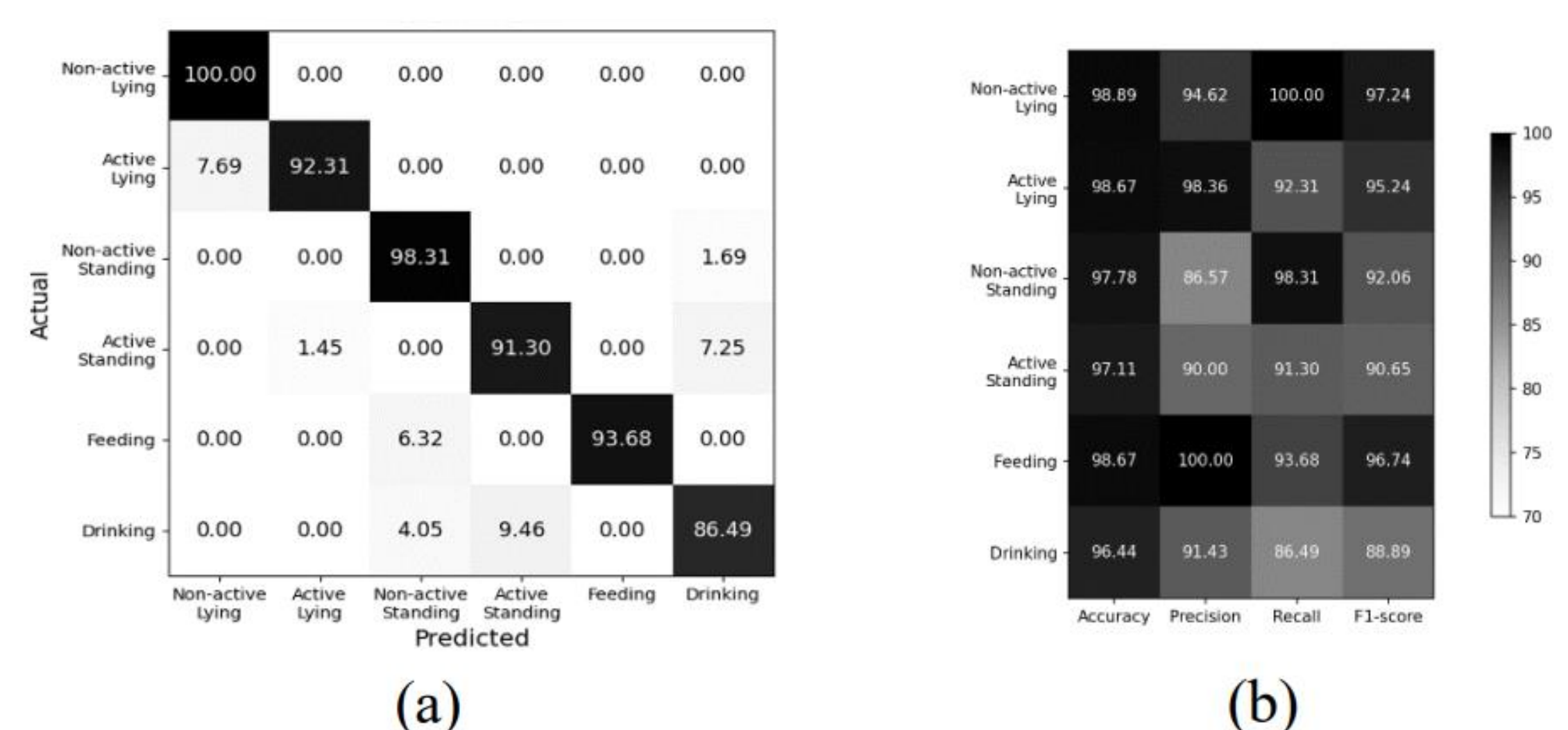


Figure 2. Evaluation of Behavior Recognition Model. (a) Confusion Matrix, (b) Performance Metrics.

### behavior analysis

- As shown in Figure 3, for most of the time, calves exhibit a predominant state of *Non-active Lying*, with a total duration of 15 hours.
- The frequent incidence of *Standing*, *Feeding*, and *Drinking* behaviors at 06:00 and 18:00 corresponds with the schedule provided by the caretaker for their feeding and cleaning operations.

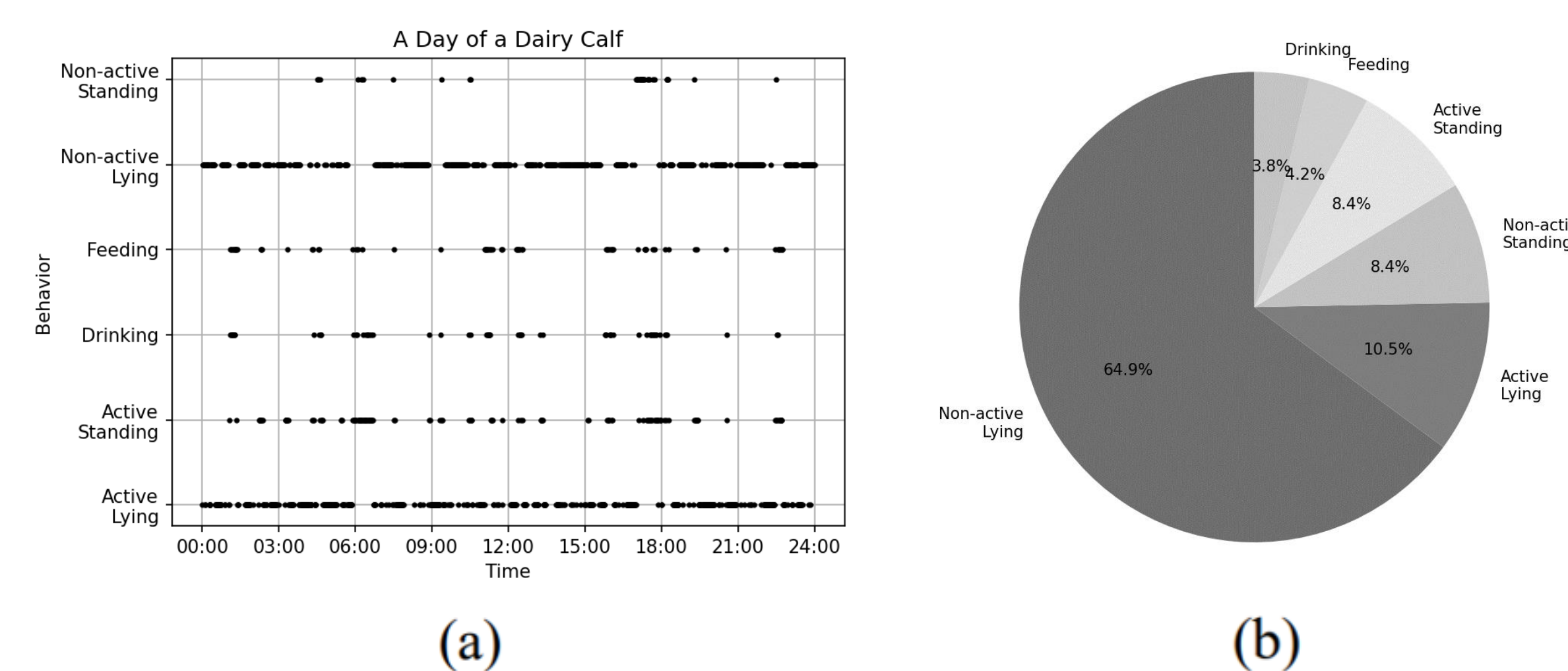


Figure 3. Behavior Analysis Results. (a) Daily Behavioral Distribution, (b) Total Duration Proportions for Each Behavior.

## Conclusion

In this study, a spatio-temporal Convolutional Neural Network was utilized for dairy calf action recognition, achieving a top-1 accuracy of 0.94 and an F1-score of 0.93. The model was then utilized to analyze behavior patterns over time, providing valuable insights for both calf behavior research and farm management decision-making.