Journal of Signal Processing, Vol.28, No.4, pp.179-182, July 2024

SELECTED PAPER AT NCSP'24

Precision Livestock Tracking: Advancements in Black Cattle Monitoring for Sustainable Agriculture

Su Myat Noe[1], Thi Thi Zin[2], Pyke Tin[2] and Ikuo Kobayashi[3]
[1]Interdisciplinary Graduate School of Agriculture and Engineering,
[2]Graduate School of Engineering,
[3]Faculty of Agriculture,
University of Miyazaki
1-1 Gakuen Kibanadai-Nishi, Miyazaki 889-2192, Japan
Email: {z319701@student, thithi@cc, ikuokob@cc}.miyazaki-u.ac.jp,
pytetin11@gmail.com

Abstract

Utilizing computer vision for animal behavior monitoring provides significant benefits by minimizing direct handling and capturing diverse traits through a single sensor. However, accurately identifying animals remains a challenge. To address this, this study introduces an innovative approach to monitor black cattle in dynamic agricultural environments to ensure their health welfare. By integrating advanced techniques like DETIC for automated labeling and YOLOv8 for real-time detection, the research emphasizes improving accuracy and robustness in tracking black cattle tracking within complex open ranch environments. Moreover, the customized ByteTrack model tailored for ranch scenarios significantly enhances cattle tracking across intricate landscapes. Achieving a mean Average Precision (mAP) of 0.901 and a Multi-Object Tracking Accuracy (MOTA) of average accuracy 92.185% of four videos, this approach appears to offer a viable resolution for conducting individual cattle behavior analysis experiments through the application of computer vision.

1. Introduction

Cattle hold immense significance in livestock farming, serving as a primary protein source across various cultures and regions. The assurance of their well-being and the optimization of productivity relies heavily on monitoring individual cattle behavior effectively. Yet, conventional manual monitoring encounters considerable hurdles due to staffing limitations and workflow constraints, rendering them impractical and unsustainable. In today's competitive landscape of modern, intelligent farming, the efficient and scalable management of cattle farms stands as a critical necessity.

To tackle this challenge, we introduce a novel solution automatic labelling technique aims for cattle tracking. The precise delineation of cattle areas forms a pivotal stage in this process. By harnessing the capabilities of the state-of-the-art Byte Track, a deep learning-based tracking framework, our method facilitates real-time, automated monitoring of cattle behavior. Embracing these sophisticated systems empowers farm managers to elevate cattle health management, optimize productivity, and enhance the overall sustainability and profitability of their operations, ushering in a new era of intelligent and efficient livestock management.

Furthermore, the evolution of automated segmentation models has triggered a significant transformation in tasks related to detecting and recognizing objects. However, prior to the emergence of these models, segmenting datasets containing objects that closely resemble the background color posed substantial challenges. Conventional image processing techniques were commonly employed to tackle this issue, yet their efficacy was limited, particularly in complex scenarios. Additionally, tracking livestock in complex environments often resulted in missed and falsely detected instances during individual tracking. Another approach involved manual annotation for instance segmentation, though it was time-consuming, labor-intensive, and not cost-effective [1].

Detector with Image Classes (DETIC) developed by Facebook Research [2], Detecting Twenty-thousand Classes using Image-level Supervision for precise automatic labelling utilized as a pre-processing method. An instance segmentation model that can identify 21K object classes, DETIC demonstrates strong zero-shot performance across various segmentation tasks. As a robust tool for image segmentation, it holds significant promise for numerous applications.

This proposed research proposed to harness the advancements in generative AI models and formulate an effective and efficient automatic segmentation approach specifically tailored for tracking black cattle. The following are the contributions:

- Employing the DETIC Model for automated cattle region labeling and implementing semi-supervised learning with the resulting output weight aims to enhance both efficiency and accuracy, thereby

This work was partially supported by JST SPRING, Grant Number WN922001.

optimizing resource utilization for improved precision within a more efficient timeframe.

- Incorporating YOLOv8-ByteTrack, an advanced multi-object tracking algorithm specifically tailored for tracking multiple black cattle in an open ranch using a single camera.
- Modifications to the YOLOv8 anchor boxes by employing the K-means clustering algorithms to detect the small cattle in the ranch.

These enhancements notably boost the network's capacity to accurately extract features from cattle targets, facilitating the capture of long-range dependencies within the tracking process. To evaluate the efficacy of the proposed approach, the rest of this paper is organized as follows: Section 2 introduces materials and methods. Section 3 shows the experimental results and finally section 4 concludes the paper.

2. Materials and Methods

2.1 Integration of DETIC model

In this section, the step-by-step procedures are described in the following Figure 1. As our datasets are mostly morning and evening data because of the nature of the open dataset, the cattle are not available on the ranch in the afternoon because of the roaming time outside the ranch. So, we need to adjust the time and make the histogram equalization to adjust the brightness as a pre-processing stage. Then, we will do the next step as an automatic labelling process with DETIC model as a focus. We will also compare other automatic segmentation models with SAM and Grounding Dino model as a performance evaluation.

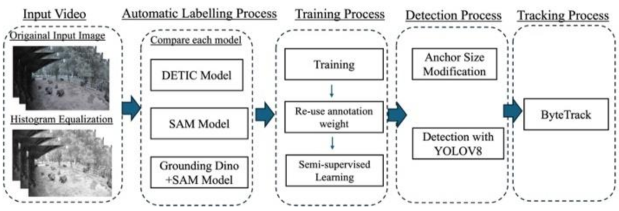

Figure 1: Overall proposed system

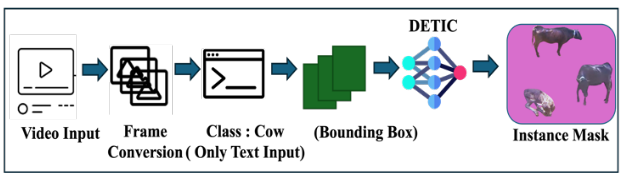

Figure 2: Step by step process flow of the automatic process with DETIC model

Our approach to achieving automatic labelling follows a systematic process. Initially, we curate a diverse range of images and apply the DETIC model to segment individual images, outlined in Figure 2. This formatting standardizes annotations, allowing seamless integration with the YOLOv8 model. Our tailored dataset specifically addresses the need for precise and efficient cattle segmentation.

2.2 Dataset partitioning and model training

We adopted 70:30 split strategies for training and validation data during the YOLOv8 [3] model's training [4]. This division aimed to provide the model with a robust dataset for learning while ensuring its capacity to generalize well to new, unseen data. In response to the challenge posed by varying illumination in our outdoor open ranch dataset, we meticulously curated morning and evening data for training. This deliberate selection allowed for a balanced representation of diverse lighting conditions, enabling the model to adeptly adapt to these variations. The training phase encompassed 3000 images, comprising 45000 instances, and persisted for 100 epochs, employing a batch size of 8. Stochastic Gradient Descent (SGD) with a momentum of 0.937 optimized the model, initialized with the YOLOv8 segmentation weight file. The model's training duration totaled 3 hours duration long.

2.3 Model validation and augmentation techniques

The YOLOv8 model represents an anchor-free approach to object detection, yet its predefined set of anchor boxes might not universally cater to all datasets. To address this, we made a customizing the anchor boxes to better suit our dataset. Employing the K-means clustering algorithm allows us to discern the prevailing object sizes and shapes within our specific dataset. This method facilitates the identification of the most common characteristics, empowering us to refine and adjust the anchor boxes accordingly. By tailoring these anchor boxes to align more closely with the inherent characteristics of our dataset, we aim to optimize the model's performance and accuracy in object detection tasks.

3. Experimental Results

3.1 Performance evaluation

The DETIC model exhibited limitations in segmentation performance. While initially detecting and segmenting a considerable number of cattle objects, it occasionally omitted some within the frame, necessitating manual re-annotation for satisfactory results. This highlights that relying solely on the DETIC model might not suffice for comprehensive and precise cattle segmentation. So, to notice the performance comparison,

Figure 3 outlines the input and output performance detail. Figure 4 describes the comparison result with the other model such as Grounding Dino + SAM, SAM models [5,6].

Figure 3: Segmentation result of the DETIC model

Figure 4: Comparison result of the automatic segmentation models

3.2 Detection evaluation

In this segment, we outline the outcomes derived from a series of YOLOv8 detection models, each tailored with distinct configurations targeting specific performance aspects. Our assessment of these models encompasses three pivotal metrics: precision, mAP, computed across an Intersection over Union (IoU) range from 0.5 to 0.95, and inference time. Figure 5 describes the comparison with other deep learning-based detection models which are: YOLOv5, YOLOv7, and Detectron 2.

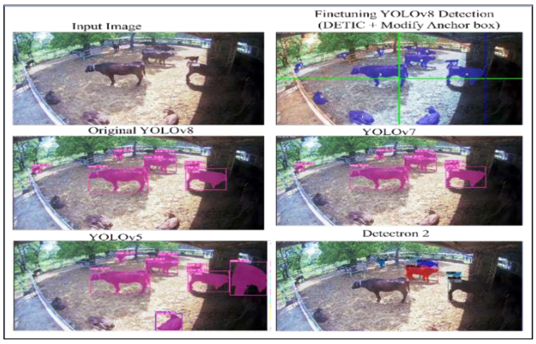

Figure 5: Comparison of detection results using different detection models

3.3 Tracking evaluation

Within our dataset comprising black cattle, we devised a specialized approach that combines the strengths of YOLOv8 for instance segmentation and Byte Track for multi-object tracking. This model effectively identifies and pinpoints individual cattle instances in each video frame, generating precise bounding box predictions along with class labels. This initial segmentation step forms the groundwork for subsequent tracking procedures. Our open ranch dataset, encompassing an area of 23.3 meters by 20 meters, serves as the home to approximately 55 cattle. Managing this substantial livestock population poses a noteworthy challenge. In response to this challenge, we have implemented the YOLOv8 combination with Byte Track tracking process within our ranch management strategy. It is worth emphasizing that our ranch operates as an open environment, allowing cattle to move freely in and out of the premises at various intervals. To facilitate efficient tracking and management, we have implemented a tracking approach utilizing Byte Track, which employs cattle ID within each zone. By leveraging the YOLOv8 method and adapting to the dynamic nature of our open ranch, we have enhanced our ability to manage our cattle effectively while ensuring their well-being. For the tracking approach involved optimizing specific parameter settings, such as a track threshold of 0.5, a maximum age of 500 frames, and an IoU threshold of 0.3. Careful selection of these configurations aimed to increase tracking performance, ensuring accurate object association, and maintaining consistent tracking across video frames. Our method's tracking results, particularly from the morning 6 am video without histogram equalization, underscore the effectiveness of the DETIC segmentation and Byte Track tracking algorithm in detecting and tracking cattle instances.

Employing the Multiple Object Tracking Accuracy (MOTA) metric in our research facilitated a comprehensive evaluation of our tracking approach's effectiveness on the black cattle dataset, as detailed in Table 1. Each of the videos (Camera 1 - 4) represents multiple cameras operating simultaneously for 30 minutes at 30 frames per second (fps). The total inference time for all videos is 43.629 minutes, with an average accuracy of 92.185%, calculated using equation (1) of the *MOTA* calculation, which considers False Negative (*FN*), False Positives (*FP*), ID switches (*IDSW*), and Ground truth detection (*gDet*) information.

$$MOTA = 1 - \frac{(|FN|+|FP|+|IDSW|}{|gDet|} \qquad (1)$$

Table 1. Comparison of tracking accuracy

Video Sequences	Video 1	Video 2	Video 3	Video 4
Accuracy (%)	94.36	90.49	91.32	92.57
Average Accuracy (%)	92.185			

Figure 6 described the outcomes results from frames 1000 to 4000 for each video. When new cattle enter the camera's field of view, we temporarily withhold identification. After a delay of 30 frames, we assigned an identification (ID). This methodical approach ensures accurate tracking, enabling robust analysis and interpretation of the results.

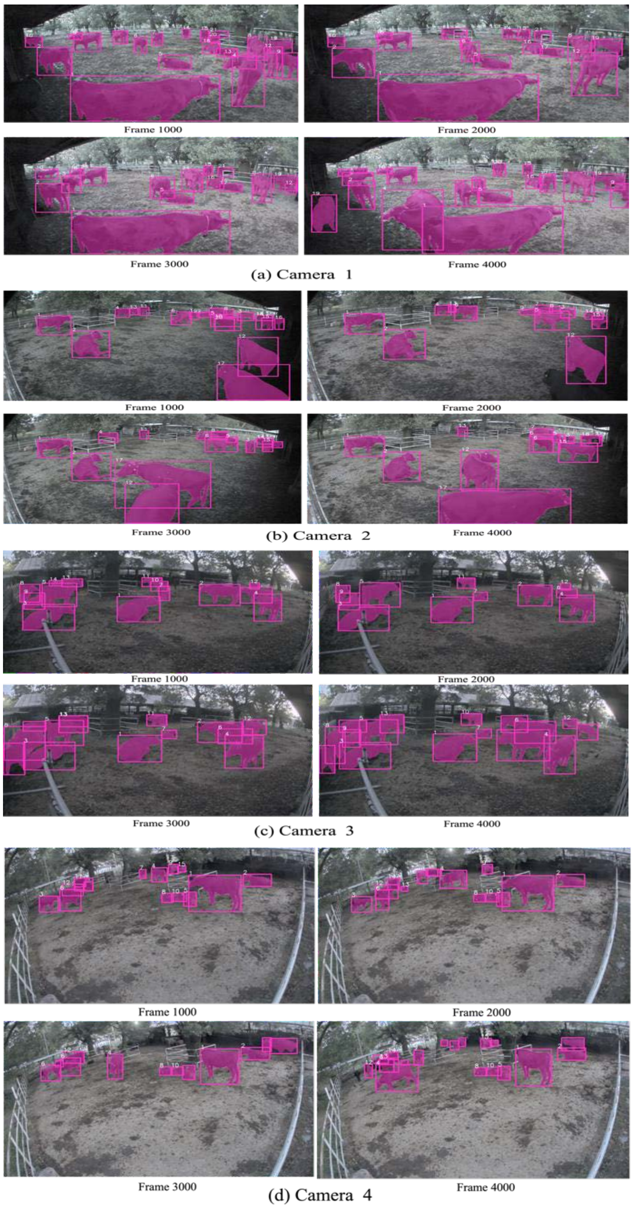

Figure 6: Result of the fusion of YOLOV8 and ByteTrack with different camera perspectives

Despite challenging low-light conditions, these models leverage contextual cues such as cattle shape and movement, enabling accurate tracking. While instances of detection failure occur in partially occluded or crowded areas, our approach demonstrates resilience in handling scenarios where cattle may be partially hidden or obscured by other objects.

4. Conclusions and Future Work

Our study presents a pioneering approach in tracking black cattle, harnessing a fusion of the DETIC model and YOLOv8 alongside the Byte Track model, effectively integrating motion and appearance details. Notably, our findings exhibit exceptional results, particularly in handling multi-object tracking challenges. The application of our approach to the black cattle dataset within open ranch settings, known for their intricate and crowded cattle scenes, underscores its practical applicability in real-world cattle monitoring scenarios.

The pivotal strength of our approach lies in its ability to furnish robust and accurate tracking, owing to the integration of contextual cues. This research sheds light on the efficacy of advanced tracking algorithms, establishing their significance in livestock management and related applications. Overall, our study offers a pragmatic and efficient solution specifically tailored for tracking black cattle, effectively tackling practical challenges, and enhancing livestock monitoring methodologies. As the livestock industry continues its embrace of cutting-edge technologies, our research marks a significant stride towards the development of more sophisticated and efficient cattle tracking methods. Ultimately, this contributes substantially to the evolution of contemporary farming practices and the promotion of sustainable livestock management.

References

[1] Su Myat Noe, Thi Thi Zin, Pyke Tin, I. Kobayashi: Comparing state-of-the-art deep learning algorithms for the automated detection and tracking of black cattle, Sensors, Vol. 23, No. 1, p. 532, 2023.

[2] X. Zhou, R. Girdhar, A. Joulin, P. Krahenbuhl and I. Misra: Detecting twenty-thousand classes using image-level supervision, European Conference on Computer Vision, pp. 350-368, 2022.

[3] J. Terven, DM. Córdova-Esparza, JA. Romero-González: A comprehensive review of yolo architectures in computer vision: from yolov1 to yolov8 and yolo-nas, Machine Learning and Knowledge Extraction, Vol. 5, No. 4, pp.1680-1716, 2023.

[4] Y. Zhang, P. Sun: Bytetrack: multi-object tracking by associating every detection box, European Conference on Computer Vision, pp. 1-21, 2022

[5] A. Kirillov, E. Mintun: Segment anything, Proceedings of the IEEE/CVF International Conference on Computer Vision, pp. 4015-4026, 2023.

[6] S. Liu, et. al.: Grounding dino: Marrying dino with grounded pre-training for open-set object detection, arXiv preprint arXiv:2303.05499, 2023.