JAST (Journal of Animal Science and Technology) TITLE PAGE
Upload this completed form to website with submission

ARTICLE INFORMATION	Fill in information in each box below
Article Type	Research article
Article Title (within 20 words without abbreviations)	Automatic Identification and Analysis of Multi-Object Cattle Rumination Based on Computer Vision
Running Title (within 10 words)	Automatic Identification and Analysis of Multi-Object Cattle Rumination
Author	Yueming Wang1, Tiantian Chen1, Baoshan Li1, Qi Li1
Affiliation	1 School of Information Engineering, Inner Mongolia University of Science and Technology, Baotou, 014010, China
ORCID (for more information, please visit https://orcid.org)	Yueming Wang (https://orcid.org/0000-0002-5810-0894) Tiantian Chen (https://orcid.org/ 0000-0003-2468-5262) Baoshan Li (https://orcid.org/0000-0001-8112-7916) Qi Li (https://orcid.org/0000-0002-9797-5159)
Competing interests	No potential conflict of interest relevant to this article was reported.
Funding sources State funding sources (grants, funding sources, equipment, and supplies). Include name and number of grant if available.	This research was funded by the Inner Mongolia Major Science and Technology Project (Grant NO.2019ZD025), the Natural Science Foundation of Inner Mongolia (Grant NO.2020MS06008), and the Key Technology Projects of Inner Mongolia Autonomous Region (Grant NO.2021GG0045).
Acknowledgements	We would like to appreciate the funding organization for their financial supports. Many thanks to the helpful comments and suggestions provided by all the authors cited in this article and the anonymous reviewers
Availability of data and material	Upon reasonable request, the datasets of this study can be available from the corresponding author.
Authors' contributions Please specify the authors' role using this form.	Conceptualization: Wang YM, Chen TT. Data curation: Wang YM, Chen TT. Formal analysis: Wang YM, Li BS. Methodology: Wang YM, Chen TT, Li Q. Software: Chen TT. Validation: Wang YM, Chen TT. Investigation: Li BS, Li Q. Writing - original draft: Wang YM, Chen TT. Writing - review & editing: Wang YM, Chen TT, Li Baoshan, Qi Li.
Ethics approval and consent to participate	This article does not require IRB/IACUC approval because the data in this research was collected by cameras, we didn't touch cow's body.

CORRESPONDING AUTHOR CONTACT INFORMATION

For the corresponding author (responsible for correspondence, proofreading, and reprints)	Fill in information in each box below
First name, middle initial, last name	Tiantian Chen
Email address – this is where your proofs will be sent	ctt010908@163.com
Secondary Email address	2020023125@stu.imust.edu.cn
Address	School of Information Engineering, Inner Mongolia University of Science and Technology, Baotou, 014010, China
Cell phone number	+86-13848014025
Office phone number	None
Fax number	None

1   **Abstract**

2   Rumination in cattle is closely related to their health, which makes the automatic monitoring of

3   rumination an important part of smart pasture operations. However, manual monitoring of cattle

4   rumination is laborious and wearable sensors are often harmful to animals. Thus, we propose a

5   computer vision-based method to automatically identify multi-object cattle rumination, and to

6   calculate the rumination time and number of chews for each cow. The heads of the cattle in the video

7   were initially tracked with a multi-object tracking algorithm, which combined the You Only Look

8   Once (YOLO) algorithm with the Kernelized Correlation Filter (KCF). Images of the head of each

9   cow were saved at a fixed size, and numbered. Then, a rumination recognition algorithm was

10   constructed with parameters obtained using the frame difference method, and rumination time and

11   number of chews were calculated. The rumination recognition algorithm was used to analyze the head

12   image of each cow to automatically detect multi-object cattle rumination. To verify the feasibility of

13   this method, the algorithm was tested on multi-object cattle rumination videos, and the results were

14   compared with the results produced by human observation. The experimental results showed that the

15   average error in rumination time was 5.902% and the average error in the number of chews was

16   8.126%. The rumination identification and calculation of rumination information only need to be

17   performed by computers automatically with no manual intervention. It could provide a new

18   contactless rumination identification method for multi-cattle, which provided technical support for

19   smart pasture.

20   **Keywords:** Cattle; Rumination; YOLOv4; KCF; Frame difference

21

1. Introduction

23   Owing to the rich grasslands in the Inner Mongolia Autonomous Region (China), animal husbandry

24   is an important part of the local economy, and its development is a leading concern in China [1].

25   Cattle breeding is a particularly important field in animal husbandry. With the expansion of breeding

26   scale and the development of information technology, cattle breeding has gradually changed from a

27   traditional cage-free model to a specialized cluster model [2]. Smart pasture could manage land and

28   animals more effectively through automatic equipment and technology, and the data of individual

29   animal could be collected, which contributes to improve farm animal welfare, reduce manpower,

30   increase productivity, and boost profits [3-4]. The accurate monitoring of the behavior of individual

31   animals plays an important role in assessing their physical health [5].

32   Rumination is a unique behavior observed in ruminants, such as cattle and sheep. Accurate

33   monitoring of rumination time and frequency can help the farmer assess the physical health of the

34   cattle, allowing sick cattle to receive treatment as soon as possible [6-8]. It can also help determine

35   whether to modify ration particle size according to the amount of time each cow spends ruminating,

36   thereby achieving precise feeding and increasing the revenue earned by herders [9, 10]. There are

37   older studies on the automatic monitoring of rumination, which are mainly divided into two

categories: one category identifies rumination by fitting animals with contacting monitoring devices [11-15], while the other monitors the animals via visual rumination monitoring programs [16-20]. These wearable devices have been used for the automatic monitoring of cattle rumination, and the accuracy of such devices have met the basic commercial requirements the accuracy. Because all of them contact the body, it is more convenient to analyze videos and monitor rumination through visual rumination monitoring programs. With the development of artificial intelligence, contactless visual inspection using visual monitoring devices has recently gained much attention.

The farm environment is complex and the number of cattle captured in surveillance videos is unclear, for the realization of cow's contactless rumination analysis, it is necessary to implement multi-object tracking using computer vision. A traditional object-tracking algorithm can be created by training a discriminative or generative model, such as Multiple Instance Learning (MIL) [21], Track Learn Detect (TLD) [20] and Discriminative Scale Space Tracker (DSST) [23]. With the ongoing development of artificial intelligence, researchers more commonly apply deep learning to multi-object tracking, as with the Hierarchical Correlation Features-based Tracker (HCFT) [24], Accurate Tracking by Overlap Maximization (ATOM) [25], DeepSORT [26] and CenterTrack [27].

Some researchers have applied multi-object tracking algorithms to animal husbandry. Sun et al proposed an algorithm for multi-object tracking loss correction based on Faster Region-based Convolutional Neural Networks (R-CNN) after observing that target-tracking frames can get lost while visually tracking pigs [28]. Zhang et al proposed a robust online method of detecting and tracking multiple pigs, which, coupled with a CNN-based detector and a correlation filter-based tracker via a novel hierarchical data association algorithm [29]. Zhang et al successfully tracked beef cattle in a real-time surveillance video by adding a long-short range context enhancement module (LSRCEM) to the You Only Look Once Version 3 (YOLOv3) algorithm and combining it with the Mudeep re-identification model. Before rumination could be identified, the cattle heads needed to be tracked. The methods described above track animals well [30].

The aforementioned object-tracking methods mainly tracked the whole body of animal, they are not suitable for head-tracking. Xu et al used RetinaNet-based detection model for the detection of multi-view cattle faces [31]. The objective of this study was to find a new way to automatically monitor and analyse cattle rumination using visual rumination monitoring programs, with no physical contact. The cattle mouths need to be detected while detected cattle heads. So we used a multi-object tracking algorithm combined with YOLO and the Kernelized Correlation Filter (KCF) for the tracking of cattle heads [32], and the mouths were detected by YOLOv4. When the cattle were ruminating, they were generally relaxed; their bodies were mostly stationary, with movement observed only in their lower jaws. Before examining the cattle for rumination, we tracked their heads via video, rather than their whole bodies. The parameters of the KCF trackers were set to achieve stable, automatic multi-object tracking of cattle heads. The rumination recognition algorithm was then constructed using the frame difference method and was used to identify rumination, as well as to calculate rumination time and number of chews for each cow. This study could provide a new no-contact method of automatically

monitoring rumination in cattle, and provide valuable technical data for abnormal behavior analysis and precision livestock farming.

2. Materials and Methods

2.1 Data collection

2.1.1 Video capture

The videos used in this experiment were captured at a cattle farm in Baotou, Inner Mongolia Autonomous Region, China, on October 11, 2020, when the temperatures were between 1 and 10 Celsius. The videos were manually filmed outside the fence with a high-definition video camera at a distance of about 4~8 m from the cattle. After screening, three stable videos showing multi-object rumination with the least amount of jitter were chosen, and these cows could be seen clearly in these videos. Each video included at least three cows and lasted 20 to 35 seconds, with a frame rate of 30 fps. Not all cattle in the videos were ruminating at the time of filming; some were turning head, lying and standing. The video statistics are shown in Table 1.

All three videos were numbered. The video time column shows the duration of each video. The cow numbers were obtained by numbering the cows during the detection stage, and were used to identify each cow. The rumination time and number of chews were obtained via human observation, and are shown in the last two columns of Table 1.

2.1.2 Creating the dataset

In any experiment where an object detection model needs to be trained, an object detection dataset should first be created. For this study, key frames were extracted from the captured videos and combined with cattle images provided by other re-searchers in the same group, resulting in a total of 1,000 images. The object detection dataset was created with the image-labelling software LabelImg [33], we drew the rectangular boxes of cattle heads and mouths in these images manually, and the corresponding label files were generated, the labeled image were shown in Figure 1. The dataset consists of the image files and their corresponding label files. The images were then randomly divided into training and testing sets at a ratio of 9:1. The YOLOv4 object detection model was trained using these images, some of which are shown in Figure 2.

2.2 Overall program for multi-object cattle rumination identification

The overall technical program for multi-object cattle rumination identification is shown in Figure 3. The main steps in identifying rumination were as follows:

(1) Training the object detection model: The YOLOv4 model was trained using the previously prepared dataset, and the object detection model was obtained.

(2) Tracking and extracting cattle heads: The head image of each cow in the video was obtained by combining the trained YOLOv4 model with the Kernelized Correlation Filter (KCF) tracking algorithm.

(3) Constructing a rumination recognition algorithm specific to cattle: The rumination recognition algorithm was constructed using previous YOLOv4 object detection results and the frame difference method.

(4) Multi-object cattle rumination identification: After head images were obtained for all cattle in the video, the rumination recognition algorithm was used to determine whether each cow was ruminating, rumination time, and number of chews.

2.3 Training the object detection model

YOLOv4 was chosen for object detection. The platform was Ubuntu 16.04.7 and the processor was an Intel(R) Core(TM) i9-9900K at 3.60GHz, with two 11GB NVIDIA GeForce GTX 2080Ti GPUs and 64GB RAM. The YOLOv4 model structure is shown in Figure 3a. Figure 4 shows some of the object detection results by the trained model.

2.4 Tracking and extracting cattle heads

Rumination is mainly a lower-jaw movement. Before rumination can be identified, head images must first be extracted. The algorithm created for this study used YOLO and KCF for multi-object cattle head tracking.

Henriques et al proposed KCF, a discriminant object-tracking algorithm [34]. A discriminant classifier was trained with a given sample to determine if the tracked object was the target. We needed to select the target area in the image to initialize the KCF tracker. A large number of positive and negative samples were generated by cyclic shift sampling around the target area, and ridge regression was used to train the tracker. For each new frame, the tracker detected the patch at the previous position, and the target position was updated to the one that yielded the maximum value. Per the diagonalization of circulant matrices in the Fourier domain, this greatly reduced computation and improved operational speed to meet the real-time requirements.

The multi-object cattle-tracking algorithm flowchart is shown in Figure 3b. The main steps were as follows:

(1) Object detection: The first video frame was detected by calling the trained YOLOv4 model. If any cattle heads were detected, all detection boxes were sorted by the size of the abscissas of the upper left points, and this information was saved. If no heads were detected, we proceeded to the next frame, then the next, until the target was detected.

(2) Create a multi-object tracker: A multi-object tracker, including multiple KCF trackers, was created. The number of KCF trackers matched the number of YOLO detection boxes. The KCF trackers were initialized according to the order of the YOLO detection boxes in step (1), then numbered.

(3) Object tracking: When the KCF tracker was used to track a target, it was unable to adapt to scale changes or fast movements, so it was likely to lose the target. Therefore, the KCF tracking boxes were updated to YOLOv4 detection boxes in this study, and the cow id should be determined. Every 30 frames, the YOLOv4 model was called again to detect the image, and the cattle head detection

boxes were obtained. The new YOLO detection boxes needed to be assigned the cow id number, and judged whether the KCF tracking boxes need to be reinitialized according to the tracking rules.

(4) Tracking rules: The specific tracking rules about tracker reinitialization could be seen in Figure 5. If the number of detected boxes was different than the number of KCF tracking boxes, the KCF trackers needed to be reinitialized. Otherwise, through calculating which KCF box had the minimum distance with the YOLO boxes, the YOLO boxes were assigned a corresponding id.

And we needed to judge whether there were repetitive id number of cows. If it was, the minimum distance of the same id number were compared, the larger one was assigned a new id, and the KCF trackers were reinitialized with the YOLO detection boxes. Otherwise, we would determine whether the distance between detection boxes and tracking boxes with the same id were all in the threshold range, as expressed as Equation 1.

$$\sqrt{(x_y - x_k)^2 + (y_y - y_k)^2} < s \tag{1}$$

(x_y, y_y) and (x_k, y_k) represent the coordinates of the YOLO detection box center point and the KCF tracking box center point respectively, and s is the threshold to determine the distance between the two center points.

If it was, the KCF trackers continued tracking the cow heads. Else, the KCF trackers were reinitialized with the YOLO detection boxes. The head images of each cow were extracted, then stored in their respective positions at a fixed size.

2.5 Constructing a rumination recognition algorithm specific to cattle

When we used video monitoring, we were able to see some cattle heads, but could not recognize any rumination processes due to object occlusion. Therefore, we needed to determine whether the rumination of the target cattle could be identified. When the trained YOLOv4 model was called, the mouths of the cattle were also detected, and the mouth detection boxes were stored. If no mouth was detected in the image, we were unable to identify rumination and therefore proceeded to the next frames until the mouth was detected. When a mouth was detected, we determined whether there was a center point in the mouth detection box. If there was a center point, we continued with the rumination identification of the target cow. If there was no center point, we determined that the rumination of the target cow could not be identified, and continued to the next.

Because any head movements were small when the cattle were ruminating, head motion contour images were extracted using the frame difference method and then transformed into binary images to reflect the size of the changes. We could see any changes in position through the number of white pixels in the binary images. The comparisons between ruminating and non-ruminating cattle can be seen in Figure 6. If a cow was ruminating, its mouth would repeatedly open and close. The head changes are shown in Figure 6a: the changes were mainly in the lower jaw, and movements at the top of the head were small. If the cow was doing something else, such as turning its head (as in Figure 6b), the changes at the top of the head and around the lower jaw were large. If the cow was static, as in

1 Figure 6c, there was almost no variation in the head images. The top of the head and the lower jaw
2 should change at the same time during rumination, unless the cattle are stationary. If the changes to
3 the top of the head during rumination were smaller than a head turn, head-turning could be excluded
4 from any changes noted to the top of the head, and rumination could be recognized.

5 Therefore, the recognition algorithm described in this paper mainly used the frame difference
6 method. After the head images of the target cattle were obtained, they were converted to grayscale,
7 and any background noise was removed using median filtering. The head motion contour images were
8 extracted using the frame difference method to differentiate adjacent frames of the same target cow's
9 head, then trans-formed into binary images. Because the main difference during rumination lies in the
10 changes to the top of the head and the lower jaw, the tops of the cows' heads and the lower jaws were
11 separated at a ratio of 2:1 to better identify rumination, as shown in Figure 7

12 In this study, the white pixels in the two parts of the binary images were counted. The head changes
13 seen in some adjacent video frames during rumination were small, so the number of white pixels in
14 these video frames might have been too small to accurately recognize rumination. The average
15 number of white pixels in continuing 30 frames showing the top of the head and the lower jaw ($avert_i$
16 and $averj_i$) were calculated as rumination parameters, ($avert_i$ and $averj_i$) were calculated as Equation 2.

17
$$\begin{cases} avert_i = \dfrac{(t_i + t_{i-1} + \cdots + t_{i-29})}{30} \\ averj_i = \dfrac{(j_i + j_{i-1} + \cdots + j_{i-29})}{30} \end{cases} \tag{2}$$

18 t_i is the number of white pixels in the top-of-the-head binary image of the first i frame , and j_i is the
19 number of white pixels in the lower jaw binary image of the first i frame.

20 Given that rumination mainly involved jaw movement, the number of white pixels at the top of the
21 head and the lower jaw changed within a certain range. Because each camera angle was different, the
22 rumination threshold interval changed for each video. About three to five seconds of each video was
23 captured to determine the threshold value. From the shortened video, the minimum of $averj_i$ was
24 obtained as minj, the maximum of $averj_i$ was obtained as maxj, the maximum of $avert_i$ was obtained
25 as maxt. When a cow was ruminating, the number of white pixels in the binary image of its jaw
26 ($averj_i$) fluctuated within a certain range. However, the top of the head showed little to no movement,
27 so the number of white pixels at the top of the head ($avert_i$) did not exceed the maximum. Therefore,
28 the rumination threshold interval of the jaw was determined by minj and maxj, and the rumination
29 threshold interval of the top of the head was determined by maxt. If the actual rumination parameters
30 of the cattle fell within the rumination threshold, as shown as Equation 3.We determined that the
31 target cow was ruminating, and set the rumination mark. Otherwise, the target cow was determined to
32 be engaging in other activities.

33
$$\begin{cases} 0.8minj < averj_i < 1.2maxj \\ avert_i < 1.2maxt \end{cases} \tag{3}$$

2.5.1 Statistics of rumination time

If the rumination mark could be detected, the total number of rumination frames would be recorded. According to the relationship between the number of video frames, frame rate and time, as seen in Equation 4, the rumination time of the target cow would be calculated, and was displayed on the image.

$$Time = \frac{Framecount}{fps} \tag{4}$$

Time is the rumination time of the target cow, Framecount is the number of video frames during which rumination takes place, and fps is the frame rate.

2.5.2 Calculating the number of chews

During rumination, the mouths of the cattle repeatedly opened and closed. The movement of the lower jaw showed some regularity; the number of white pixels in the binary images changed consistently, as seen in Figure 8. The abscissa represented the number of video frames; the ordinate represented the number of white pixels in the lower jaw binary image. When the mouth opened, the number of white pixels gradually increased to the maximum, then reduced when the mouth closed.

The number of chews during rumination was calculated through the number of local maximum points in the change curve. Generally, cattle take more than 0.8s for every chew during rumination, and the usual video frame rate is 30. In order to reduce the errors introduced by camera jitter, the difference between the frame numbers for the local maximum point from the adjacent chewing frames should be greater than 24. The rumination time between two adjacent chewing frames should be greater than 0.8s. The number of chews for the target cattle was calculated using these rules, and was displayed above the target cattle images.

2.6 Multi-object cattle rumination identification

Head images were obtained for all cattle using the above tracking algorithm. These images were then analysed by the above rumination recognition algorithm, which identified which cattle were ruminating. The rumination time and number of chews would be calculated and displayed over the corresponding head image.

3. Results and discussion

3.1 Training result

The YOLOv4 was trained to detect the head and mouth of cattle in video. The image dataset set 608*608 pixels was input, with the max batches of 5000, learning rate of 1.5×10^{-3}, decay of 5×10^{-5}, momentum of 0.949. The loss curve of training was shown in Figure 9a, with the increasing of training batches, the loss reduced and become stable gradually. And the Precision-Recall curve of trained YOLOv4 model was shown in Figure 9b, the high area under the curve represents both high recall and high precision, where high precision relates to a low false positive rate, and high recall relates to a low false negative rate. The Precision-Recall curve shows that the trained model is

returning accurate results, as well as returning a majority of all positive results. After training, the head's AP is 98.17%, the mouth's AP is 93.60%, and mAP@0.50 is 95.9%. The object detection result was shown in Figure 10, in the condition of dark light such as the second image in Figure 10, the object could been still detected.

In previous researches, the mouth's AP of cattle was 87.8% [19], our mouth's AP was 93.6%, there is a little improvement, and the head's AP of cattle was 99.8% [31], our head's AP was 98.7%, which was a little lower than it, but according to the object detection result, the multi-object cattle heads and mouths could be detected well. And all tracking videos were converted into images, we checked them manually and couldn't find the missing detected objects, so we thought the YOLO and KCF could be used in tracking heads of multi-object cattle.

3.2 Results of rumination identification

The platform used for the identification and analysis of multi-object cattle rumination was Windows 10. The processor was an Intel(R) Core(TM) i7-7500U at 2.70GHz, with a 2GB Intel(R) HD Graphics 620 GPU and a 2GB NVIDIA GeForce 920MX GPU, and 8GB RAM.

The rumination detection result of rumination was shown in Figure 11. Green boxes indicated that the cow was ruminating; red boxes indicated that the cow was doing something else, or that rumination could not be identified. In Figure 11, the id represents the number assigned to the cow in this video, time represents total time spent ruminating, and chew represents number of chews.

The error in this study was obtained by comparing experimental results with human observation, as shown in Table 2. The error in rumination time and the error in number of chews were calculated as follows:

$$error\ in\ rumination\ time = \frac{|Stime - Atime|}{Vtime} \times 100\% \qquad (5)$$

$$error\ in\ number\ of\ chews = \frac{|Sfre - Afre|}{Vtime} \times 100\% \qquad (6)$$

Stime is the rumination time determined through our algorithm, Atime is the rumination time determined through human observation, Vtime is video duration, Sfre is the number of chews determined through our algorithm, and Afre is the number of chews determined through human observation.

Although the cattle showed small head movements, such as raising or turning their heads, the top of head movement were larger than rumination, which would resulted in more white pixels, so the rumination could still be clearly identified. The test result shows that the overall average error in rumination time was 5.902%. The maximum error in rumination time was 19.048%, and there were three cows whose rumination time errors exceeded 10%. The following were three main reasons about large error: ① the fog produced by the cow's breaths during cold weather; ② the tonguing behaviour was misidentified as ruminating; ③ the cameras sometimes took slightly shaking. In the future, in

order to improve the recognition accuracy of cattle rumination, we can try to use deep neural network to solve the problems.

The overall average error in number of chews was 8.126%. The maximum error was 15.625%, and there were three cows whose errors exceeded 10%. And the primary chewing error factors were that the previous rumination identification error was large.

Compared with the previous visual studies [16, 17, 20], the accuracy of this paper were 91.874%, we could see a slightly drop of accuracy, but the rumination of multi-object cattle could be identified basically without manual operation. The proposed method can achieve end-to-end automatic rumination identification of cattle. Although there were some errors in the experiment, if the rumination time and number of chews were detected to be abnormal frequently, the videos would be stored, and the farmers would be received a reminder and check the health status of cattle, and sick cattle would receive treatment promptly. The algorithm could save lots of labor costs and provide some technical reference for the detection of abnormal behavior and remote diagnosis in smart pasture operations.

4. Conclusions

Developing a visual rumination monitoring program is significant to the realization of smart pasture, which could identify cattle rumination automatically with no contact, and reduce the damage of wearable devices for cattle. So the paper constructed a no-contact rumination identification algorithm, the rumination time and number of chews were calculated. The average error in rumination time was 5.902% and the average error in number of chews was 8.126%. The results indicated that the proposed algorithm could be used for monitoring rumination. But the identities of individual cattle might be not determined if the tracking objects were missed. In the future, the cattle identities could be determined with other technologies, such as cattle face detection, cattle texture detection and cattle re-identification. Additionally, the rumination time and number of chews for each cow could be stored individually, and we could analyze the cattle abnormality with them.

Acknowledgments

References

1. Xie H, Wang GX, Yang YF. Analysis on the Competitiveness of Inner Mongolia Animal Husbandry Industry. Journal of Domestic Animal Ecology. 2019;40(2):50-55. https://kns.cnki.net/kcms/detail/detail.aspx?FileName=JCST201902010&DbName=CJFQ2019

2. Wang L, Tian GQ, Zhang B. Connotation, Development and Problems of China's Animal Husbandry Industry System. Chinese Journal of Animal Science. 2017;53(12):130-134. http://doi.org/10.19556/j.0258-7033.2017-12-130

3. Wolfert S, Ge L, Verdouw C, Bogaardt MJ. Big data in smart farming: A review. Agricultural Systems. 2017;153:69-80. https://doi.org/10.1016/j.agsy.2017.01.023

4. Eastwood C, Klerkx L, Ayre M, Dela Rue B. Managing Socio-Ethical Challenges in the Development of Smart Farming: From a Fragmented to a Comprehensive Approach for Responsible Research and Innovation. J Agric Environ Ethics. 2019;32: 741–768. https://doi.org/10.1007/s10806-017-9704-5

5. Chen CX, Zhang HF, Wang ZY, Wang YD. Study and application of early warning architecture of animal health culture. Transactions of the Chinese Society of Agricultural Engineering. 2010;26(11):215-220. http://doi.org/10.3969/j.issn.1002-6819.2010.11.037

6. Pahl C, Hartung E, Mahlkow-Nerge K, Haeussermann A. Feeding characteristics and rumination time of dairy cows around estrus. J Dairy Sci. 2015; 98(1):148-154. http://doi.org/10.3168/jds.2014-8025

7. Moretti R, de Rezende MPG, Biffani S, Bozzi R. Heritability and genetic correlations between rumination time and production traits in Holstein dairy cows during different lactation phases. J Anim Breed Genet. 2018;135(4):293-299. http://doi.org/10.1111/jbg.12346

8. Andreen DM, Haan MM, Dechow CD, Harvatine KJ. Relationships between milk fat and rumination time recorded by commercial rumination sensing systems. J Dairy Sci. 2020;103(9):8094-8104. http://doi.org/10.3168/jds.2019-17900

9. Meng R, Cui XD, Yu LG, Ding LY, Gao RH, Zhu J, et al. The development status and prospects of livestock and poultry precise feeding management technologies. Acta Ecologiae Animalis Domastici. 2021;42(2):1-7. http://doi.org/10.3969/j.issn.1673-1182.2021.02.001

10. Xiong B, Luo Q, Yang L. Studies on Key Thing Internet Technology for Precise Livestock Feeding. Journal of Agricultural Science and Technology, 2011;13(5):19-25. https://kns.cnki.net/kcms/detail/detail.aspx?FileName=NKDB201105006&DbName=CJFQ2011

11. Hamilton AW, Davison C, Tachtatzis C, Andonovic I, Michie C, Ferguson HJ, et al. Identification of the Rumination in Cattle Using Support Vector Machines with Motion-Sensitive Bolus Sensors. Sensors. 2019;19(5):1165. https://doi.org/10.3390/s19051165

12. Campos DP, Abatti PJ, Bertotti FL, Hill JAG, Silveira ALF. Surface electromyography segmentation and feature extraction for ingestive behavior recognition in ruminants. Comput Electron Agr. 2018;153:325-333. https://doi.org/10.1016/j.compag.2018.08.033

13. Ren XH, Liu G, Zhang M, Si YS, Zhang XY, Ma L. Dairy cattle's behavior recognition method based on support vector machine classification model. Transactions of the Chinese Society for Agricultural Machinery. 2019;50 Suppl 1:290-296. https://kns.cnki.net/kcms/detail/detail.aspx?FileName=NYJX2019S1045&DbName=CJFQ2019

14. Rombach M, Münger A, Niederhauser J, Südekum KH, Schori F. Evaluation and validation of an automatic jaw movement recorder (RumiWatch) for ingestive and rumination behaviors of dairy cows during grazing and supplementation. J Dairy Sci. 2018;101(3):2463-2475. http://doi.org/10.3168/jds.2016-12305

15. Shen WZ, Zhang AJ, Zhang Y, Wei XL, Sun J. Rumination recognition method of dairy cows based on the change of noseband pressure. Information Processing in Agriculture. 2020;7(4):479-490. https://doi.org/10.1016/j.inpa.2020.01.005

16. Chen YJ, He DJ, Fu YX, Song HB. Intelligent monitoring method of cow ruminant behavior based on video analysis technology. Int J Agr Biol Eng. 2017;10(5):194-202. http://doi.org/10.25165/j.ijabe.20171005.3117

17. Chen YJ, He DJ, Song HB. Automatic monitoring method of cow ruminant behavior based on spatio-temporal context learning, Int J Agr Biol Eng. 2018;11(4):179-185. http://doi.org/10.25165/j.ijabe.20181104.3509.

18. Song HB, Niu MT, Ji CH, Li ZY, Zhu QM. Monitoring of multi-target cow ruminant behavior based on video analysis technology. Transactions of the Chinese Society of Agricultural Engineering. 2018;34(18): 211-218. http://doi.org/10.11975/j.issn.1002-6819.2018.18.026

19. Mao YR, He DJ, Song HB. Automatic detection of ruminant cows' mouth area during rumination based on machine vision and video analysis technology. Int J Agr Biol Eng. 2019;12(1):186–191. http://doi.org/10.25165/j.ijabe.20191201.4268

20. Wu DH, Wang YF, Han MX, Song L, Shang YY, Zhang XY, et al. Using a CNN-LSTM for basic behaviors detection of a single dairy cow in a complex environment. Comput Electron Agr. 2021;182: 106016. http://doi.org/10.1016/j.compag.2021.106016

21. Babenko B, Yang MH, Belongie S. Robust Object Tracking with Online Multiple Instance Learning. IEEE T Pattern Anal. 2011;33(8):1619–1632. http://doi.org/10.1109/TPAMI.2010.226

22. Kalal Z, Mikolajczyk K, Matas J. Tracking-Learning-Detection. IEEE T Pattern Anal. 2012;34(7):1409-1422. http://doi.org/10.1109/TPAMI.2011.239

23. Danelljan M, Häger G, Khan FS, Felsberg M. Discriminative Scale Space Tracking. IEEE T Pattern Anal. 2017;39(8):1561-1575. http://doi.org/10.1109/TPAMI.2016.2609928

24. Ma C, Huang JB, Yang X, Yang MH. Hierarchical Convolutional Features for Visual Tracking. In: proceeding of 2015 IEEE International Conference on Computer Vision (ICCV); 2015; Santiago, Chile. http://doi.org/10.1109/iccv.2015.352

25. Danelljan M, Bhat G, Khan FS, Felsberg M. ATOM: Accurate Tracking by Overlap Maximization. In: proceeding of 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR); 2019; Long Beach, CA, USA. http://doi.org/10.1109/CVPR.2019.00479

26. Wojke N, Bewley A, Paulus D. Simple online and realtime tracking with a deep association metric. In: proceeding of 2017 IEEE International Conference on Image Processing (ICIP); 2017; Beijing, China. http://doi.org/10.1109/ICIP.2017.8296962

27. Zhou X, Koltun V, Krähenbühl P. Tracking Objects as Points. In: proceeding of 16th European Conference on Computer Vision (ECCV); 2020; Glasgow, UK. https://doi.org/10.1007/978-3-030-58548-8_28

28. Sun LQ, Zou YB, Li Y, Cai ZD, Li Y, Luo B, et al. Multi target pigs tracking loss correction algorithm based on Faster R-CNN. Int J Agr Biol Eng. 2018;11(5):192-197. http://doi.org/10.25165/j.ijabe.20181105.4232

29. Zhang L, Gray H, Ye X, Collins L, Allinson N. Automatic Individual Pig Detection and Tracking in Pig Farms. Sensors. 2019;19(5):1188. http://doi.org/10.3390/s19051188

30. Zhang HM, Wang R, Dong PJ, Sun HG, Li SQ, Wang HY. Beef Cattle Multi-target Tracking Based on DeepSORT Algorithm. Transactions of the Chinese Society for Agricultural Machinery. 2021;52(4):248-256. http://doi.org/10.6041/j.issn.1000-1298.2021.04.026

31. Xu B, Wang W, Guo L, Chen G, Wang Y, et al. Evaluation of Deep Learning for Automatic Multi-View Face Detection in Cattle. Agriculture. 2021;11(11):1062. https://doi.org/10.3390/agriculture11111062

32. Yuan ZF. Human Body Tracking Method Based on Deep Learning Object Detection. In: proceeding of the 2nd International Conference on Computer Science and Software Engineering (CSSE), 2019; Xi'an, China. https://doi.org/10.1145/3339363.3339390

33. Tzutalin. LabelImg. Git code (2015). https://github.com/tzutalin/labelImg

34. Henriques JF, Caseiro R, Martins P, Batista J. High-Speed Tracking with Kernelized Correlation Filters. IEEE T Pattern Anal. 2015;37(3):583-596. http://doi.org/10.1109/TPAMI.2014.2345390

1 Tables and Figures

2

Table 1. Rumination video statistics

Video number	Video time [s]	Cow numbers	Rumination time, observed manually [s]	Number of chews, observed manually
1	32	1	32	41
		2	0	0
		3	0	0
2	21	1	21	26
		2	0	0
		3	0	0
		4	21	27
		5	0	0
3	26	1	0	0
		2	23	27
		3	0	0
		4	0	0

3
4

Table 2. Experimental results and errors in this study

Video number	Cow numbers	Rumination time in our algorithm [s]	Number of chews in our algorithm	Error in rumination time	Error in number of chews	Average error in rumination time	Average error in number of chews
1	1	31	38	3.125%	9.375%	6.25%	10.417%
	2	1	2	3.125%	6.25%		
	3	4	5	12.5%	15.625%		
2	1	20	25	4.762%	4.762%	8.572%	9.524%
	2	1	2	4.762%	9.524%		
	3	3	4	14.286%	19.048%		
	4	17	21	19.048%	14.286%		
	5	0	0	0	0		
3	1	1	2	3.846%	7.692%	2.885%	4.438%
	2	21	26	7.692%	3.846%		
	3	0	0	0	0		
	4	0	0	0	0		

5
6

1
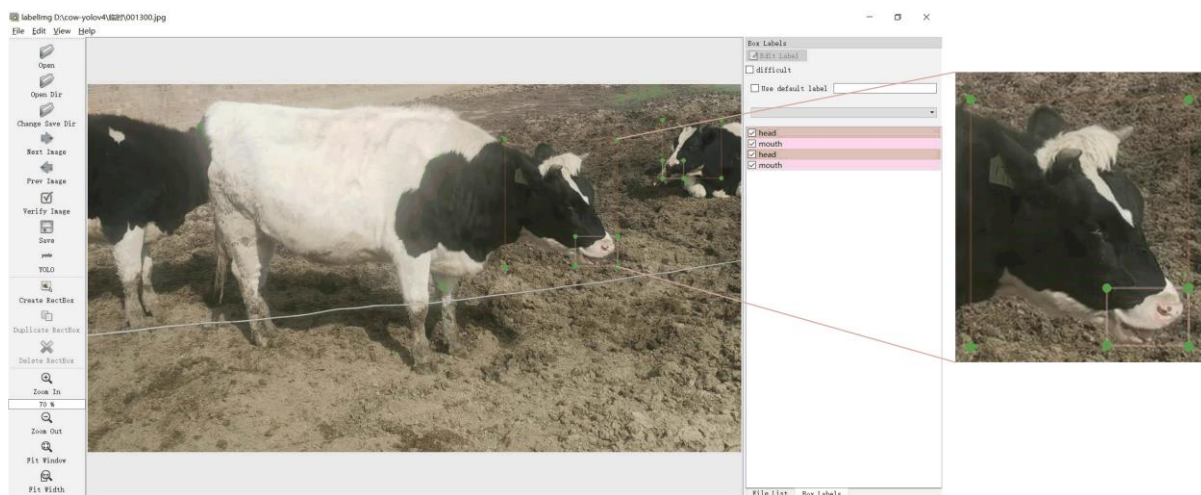
2   Figure 1. The view of labelling cattle image

3

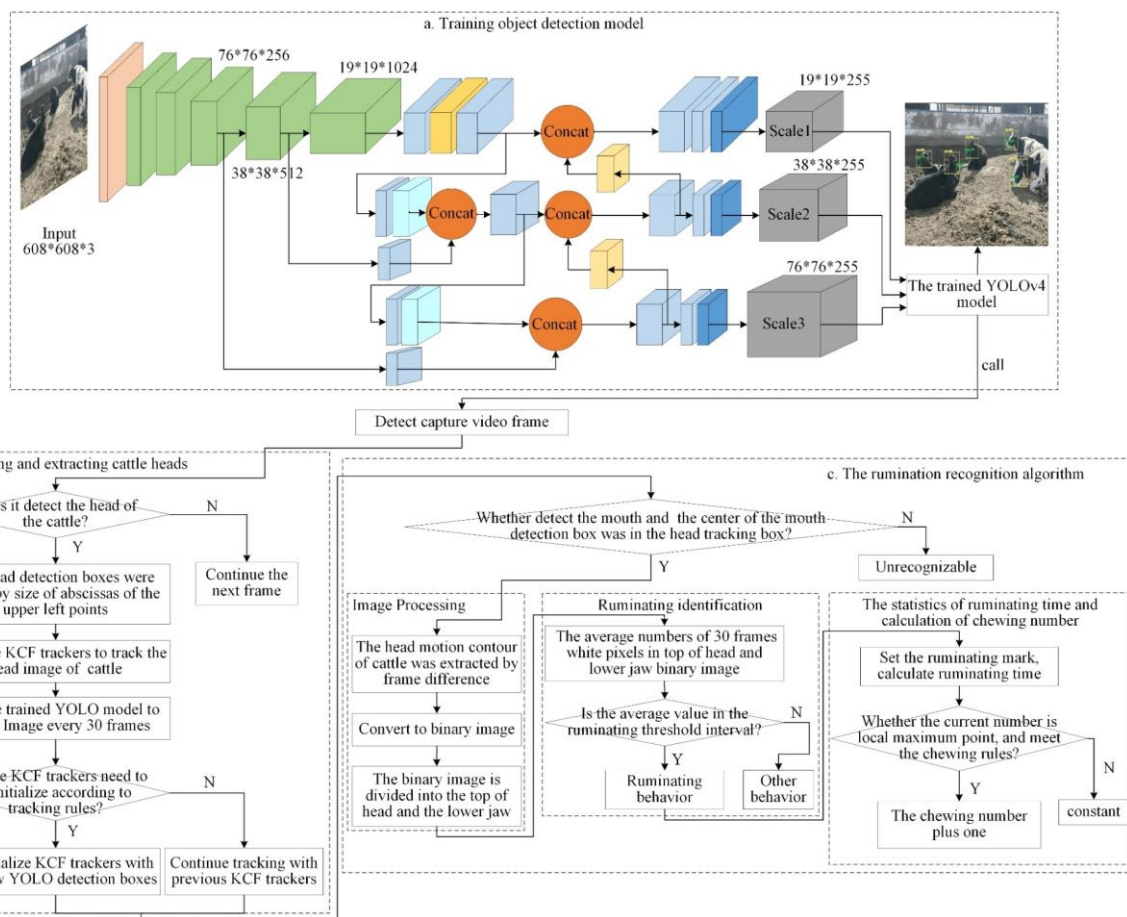4   Figure 2. Part of the training dataset

a. Training object detection model

Input
608*608*3

76*76*256

19*19*1024

38*38*512

Concat

Concat

Concat

Scale1 19*19*255

Scale2 38*38*255

Scale3 76*76*255

The trained YOLOv4 model

call

Detect capture video frame

b. Tracking and extracting cattle heads

Does it detect the head of the cattle?

N → Continue the next frame

Y

The head detection boxes were sorted by size of abscissas of the upper left points

Initialize KCF trackers to track the head image of cattle

Call the trained YOLO model to detect Image every 30 frames

Do the KCF trackers need to reinitialize according to tracking rules?

N → Continue tracking with previous KCF trackers

Y

Reinitialize KCF trackers with the new YOLO detection boxes

c. The rumination recognition algorithm

Whether detect the mouth and the center of the mouth detection box was in the head tracking box?

N → Unrecognizable

Y

Image Processing

The head motion contour of cattle was extracted by frame difference

Convert to binary image

The binary image is divided into the top of head and the lower jaw

Ruminating identification

The average numbers of 30 frames white pixels in top of head and lower jaw binary image

Is the average value in the ruminating threshold interval?

Y → Ruminating behavior

N → Other behavior

The statistics of ruminating time and calculation of chewing number

Set the ruminating mark, calculate ruminating time

Whether the current number is local maximum point, and meet the chewing rules?

Y → The chewing number plus one

N → constant

Figure 3. Schematic diagram of the overall technical solution

The detection box of head

head

mouth

The detection box of mouth

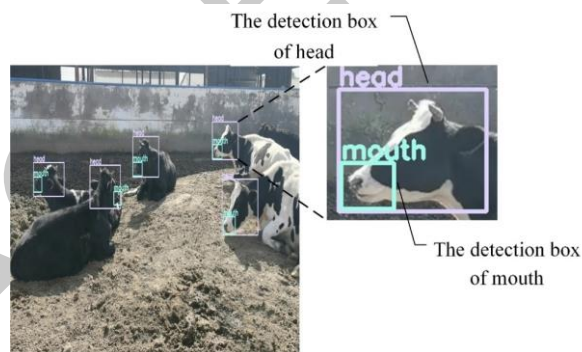Figure 4. A portion of the object detection results

1

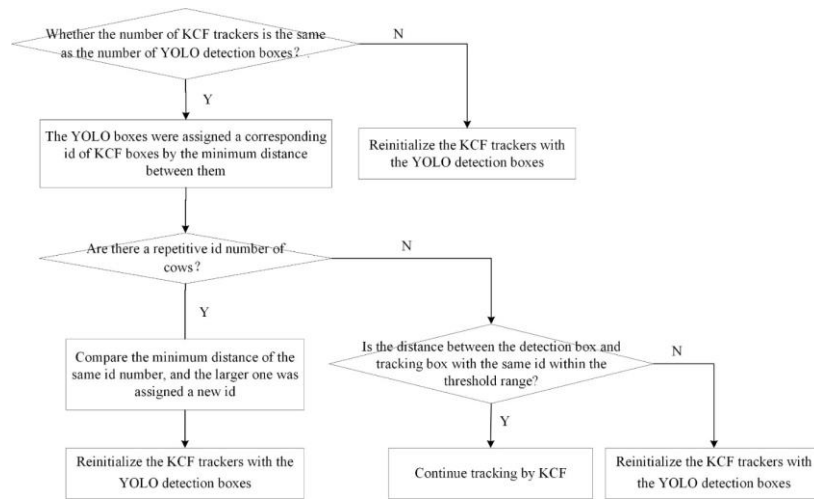2   Figure 5. The specific tracking rules

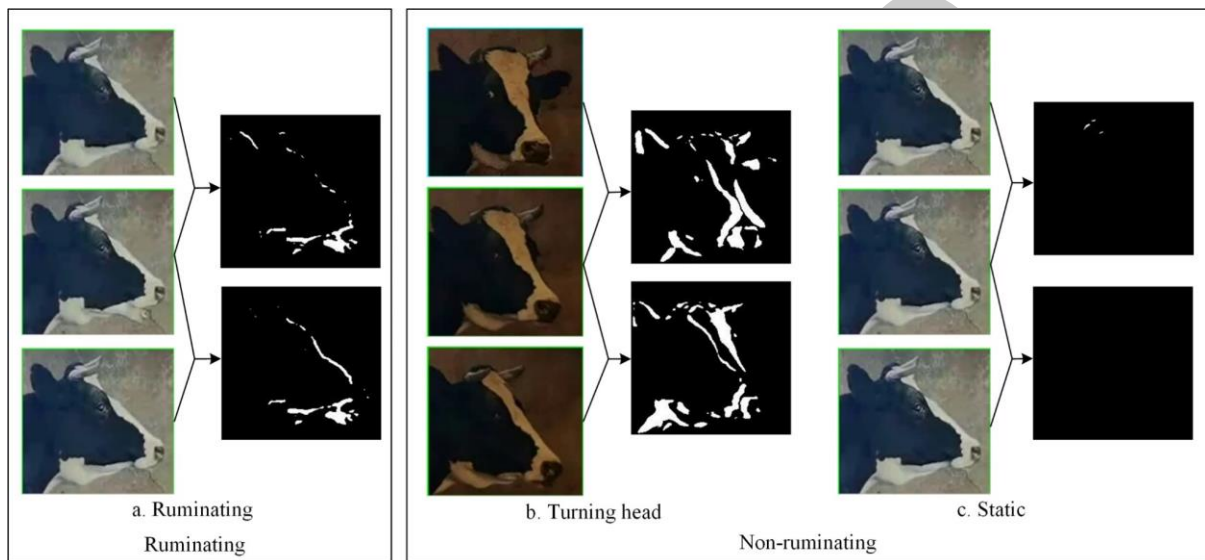
3

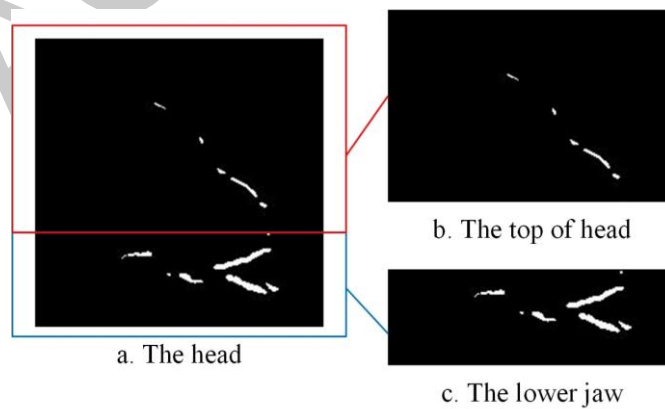4   Figure 6. Comparisons between ruminating and non-ruminating cattle

5

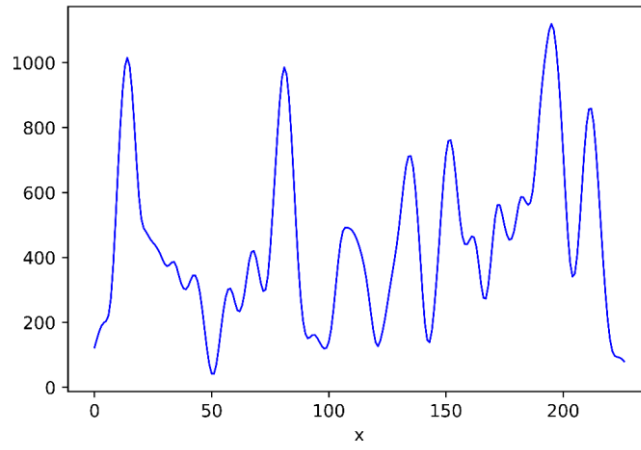6   Figure 7. Binary images of cattle head motion contours

17

1

2   Figure 8. Changes in the number of white pixels in the jaw movement contour images of cattle during
3   rumination

(a)   (b)
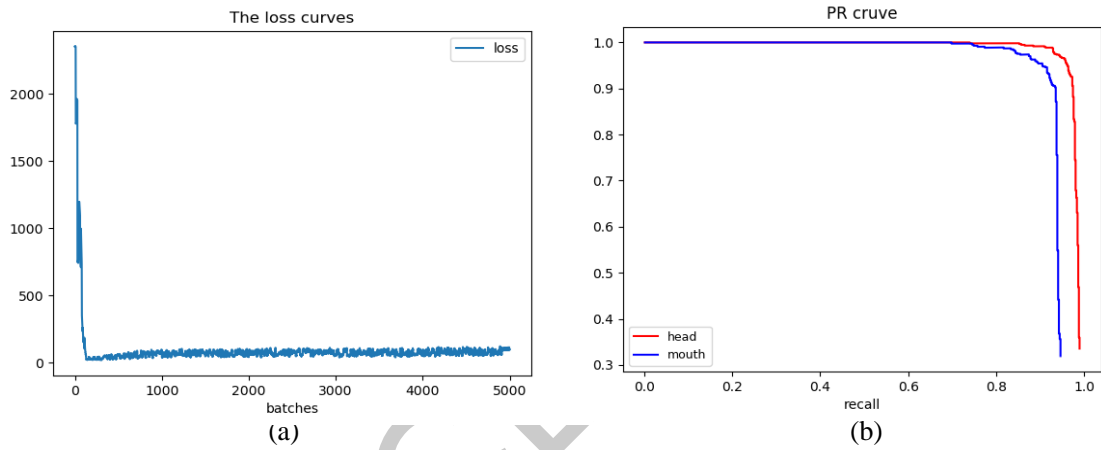
4

5   Figure 9. The training result of YOLOv4: (a) loss curve; (b) Precision-Recall curve

6

7   Figure 10. The object detection result

1   Figure 11. Video frame rumination detection results

19