Smart Agricultural Technology 9 (2024) 100566

Contents lists available at ScienceDirect

Smart Agricultural Technology

journal homepage: www.journals.elsevier.com/smart-agricultural-technology

Tracking and automatic behavioral analysis of group-housed pigs based on YOLOX+BoT-SORT-slim

Shuqin Tu [a], Yuefei Cao [a], Yun Liang [a,*], Zhixiong Zeng [b,*], Haoxuan Ou [a], Jiaying Du [a], Weidian Chen [a]

[a] College of Mathematics and Informatics, South China Agricultural University, Guangzhou 510642, China
[b] College of Engineering, South China Agricultural University, Guangzhou 510642, China

ARTICLE INFO

Keywords:
Lightweight algorithm
Behavioral analysis algorithm
MOT
1-hour video tracking

ABSTRACT

Automatic analysis of pig behavior is crucial for assessing their health and welfare status in group-housed pig farms. Currently, computer vision technology has been widely applied to recognize and track pig behavior. However, these applications lack the capability for automatic behavior analysis. Additionally, execution speed is crucial for long-term tracking, as the tracking target can easily be lost due to the complex background of pig farms, uneven lighting, and the similar appearance of pigs. To tackle these challenges, we propose a lightweight multi-object tracking (MOT) and behavioral analysis method: YOLOX+BoT-SORT-Slim. Firstly, the YOLOX detector detects pig targets and recognizes the pigs' four behaviors including "stand", "lie", "eat", and "other". Secondly, the BoT-SORT-Slim algorithm tracks the detected pigs and their behaviors at high speed. Finally, we devise a behavior automatic analysis algorithm that integrates the behavioral information and the tracking results to complete pigs' behavior analysis. To evaluate the proposed method, we conducted experiments on 1-minute, 10-minute, and 1-hour datasets. The experimental results show that on the 1-minute public and private test sets, the proposed method achieves 6.4 and 3.6 times the frame rate (FPS) of the original algorithm, with 68.05 FPS and 71.27 FPS, respectively. It also achieves a higher-order tracking accuracy (HOTA) of 84.7 % and 78.7 %, multi-object tracking accuracy (MOTA) of 98.3 % and 97.0 %, and an identification F1 score (IDF1) of 98.4 % and 93.4 %. The proposed method obtains the best results on all test sets compared to five other MOT algorithms including DeepSORT, StrongSORT, ByteTrack, OC-SORT, and BoT-SORT. In the 10-minute and 1-hour test sets, the proposed method achieves 70.28 and 72.24, 63.3 % and 76.4 %, 94.1 % and 99.6 %, and 72.5 % and 90.2 % for FPS, HOTA, MOTA, and IDF1, respectively. In terms of behavior analysis, three types of pigs' behavior analysis are plotted for each test video to assess their health status. The experimental results indicate that the proposed method performs excellently in behavior analysis and tracking, which can reliably monitor and manage pig behavior in real time, providing automated management technical support.

1. Introduction

With the advancement of artificial intelligence (AI), the modern pig industry is transitioning from household production to intensive, intelligent large-scale farming[1]. Timely diagnosis of pig health is crucial in large-scale pig farming[2,3] and observing the behavior of group-housed pigs is a primary method for assessing their well-being. In traditional pig farms, behavior monitoring of group-housed pigs relies on manual observation by stockpersons, who record behavior categories and occurrence times for target pigs. However, this monitoring method is subject to human subjectivity, requires significant manpower, and is inefficient[4]. Continuous and rapid monitoring of pig behavior using sensors, video surveillance hardware, and intelligent algorithms can detect abnormal pig activities early and enable prompt intervention. This approach is essential for improving pig welfare and overall quality of life.

Researchers have recently explored methods for pig behavior recognition using computer vision techniques. For example, Hao et al. [5]. introduced a novel deep mutual learning enhanced two-stream pig behavior recognition approach, which aims to identify behaviors such as fighting and investigating among pigs, achieving a recognition accuracy of 96.52 %. Gao et al. [1]. proposed a hybrid model that combines

* Corresponding authors.
E-mail addresses: Liang_Yun168@163.com (Y. Liang), zhixzeng@scau.edu.cn (Z. Zeng).

https://doi.org/10.1016/j.atech.2024.100566
Received 9 July 2024; Received in revised form 1 September 2024; Accepted 3 September 2024
Available online 8 September 2024

convolutional neural network (CNN) and gated recurrent unit (GRU) to differentiate aggressive and other behaviors from surveillance videos, reaching recognition accuracy of 96.52 %. Yang et al. [6]. developed a CNN-based method called Detect and Classify Posture R-CNN (DCP R-CNN) for sow detection and posture classification, which achieved a classification accuracy of 97.71 % for posture changes. Gan et al. [7]. proposed an action density estimation network (ADEN) for counting the events of piglet suckling followed by automated analysis of suckling behavior, with an average absolute error (MAE) of 0.16 %. Nie et al. [8]. integrated the ASPP module and the CReToNeXt module into the YOLOv5 model to recognize heat stress expressions in pigs, the mean Average Precision(mAP) reached 93.2%. Zhuang et al. [9]. utilized three CNN algorithms—VGG19, Xception, and MobileNetV2—to develop a recognition model for pig feeding and drinking behaviors. The MAE for feeding and drinking behaviors were 0.21 % and 0.12 % respectively. Ji et al. [10]. utilized the improved YOLOX for pig recognition, which achieved a 90.9% accuracy in seated pig recognition and an average precision of 95.7% in overall posture recognition. Notably, all the mentioned research studies relied on static images for pig behavior analysis and did not incorporate tracking techniques.

Multi-object tracking (MOT) aims to accurately locate and track multiple targets (such as animals, humans, etc.) in videos, and it has been widely applied in the field of animal husbandry. e.g., Tu et al. [11]. proposed a pig behavior tracking algorithm based on YOLOv5 fusion Byte (YOLOv5-Byte) to accurately monitor and track individual pig behavior, achieving a HOTA score of 76.5 %. Wei et al. [12]. introduced an EMA-YOLOv8 model with the ByteTrack algorithm to detect and identify the movement duration and aggressive behavior of pigs, achieving a behavior recognition accuracy of 96.4 %. Melfsen et al. [13]. employed the ByteTrack algorithm to recognize and track nine common daily behaviors of group-housed pigs, including standing, lying, playing, etc. The accuracy of the algorithm for behavioral recognition ranged from 29.6 % to 99.7 %, and its HOTA score was 98.2 %. Tu et al. [14]. used the improved DeepSORT algorithm to track the standing, lying, eating, and other behaviors of pigs with a MOTA of 98.6 %. Gan et al. [15]. developed online piglet-tracking algorithms based on graph convolutional networks (GCNs) for characterizing the social behavior of tracked preweaning piglets, achieving high precision (96.6 %) in detecting piglet social behaviors. Wang et al. [16]. adopted a one-shot tracker in which the detection and re-identification (re-ID) branches were jointly trained for tracking pigs. Their method achieved a MOTA of 91.41 % for a 2-minute video and a tracking percentage of 16.78 % per individual in an 85-minute video.

The aforementioned studies demonstrate the successful application of MOT algorithms in the short-term recognition and analysis of various pig behaviors. However, there remains limited research on long-term pig tracking in video, and the performance and execution speed in this area are not yet satisfactory[17]. Additionally, there is a lack of research on the automated analysis of behaviors. Therefore, investigating pig behavior over extended periods and the automated analysis of pig behavior under video surveillance holds significant importance for advancing our understanding of animal behavior and welfare.

To address this gap, this study proposes a lightweight algorithm for multi-object tracking and behavior analysis of pigs based on YOLOX+BoT-SORT-Slim. Firstly, the YOLOX detector detects pig targets and recognizes four behaviors ("stand", "lie", "eat", and "other"). Then, the BoT-SORT-Slim algorithm is employed for multi-object behavior tracking of group-housed pigs. Finally, based on the tracking and behavioral results, an algorithm for behavior statistics of group-housed pigs is designed to obtain individual and collective behavior analysis results. Furthermore, we validate the effectiveness of our method using both public and private datasets, comprising a total of 39 video segments. This dataset includes 34 segments of 1 min, 4 segments of 10 min, and 1 segment of 1 hour, enabling a comprehensive evaluation of our proposed approach.

This work mainly contributes to three areas, which are as follows:

(1) A lightweight multi-object tracking algorithm, YOLOX+BoT-SORT-Slim, is proposed for pig detection and behavior tracking.
(2) An analytical algorithm for group-housed pig behavior is proposed. The method completes individual and collective pig behavior analysis based on detection and tracking results.
(3) A total of 39 videos, including 34 1-minute videos, 4 10-minute videos, and 1 1-hour video, are constructed from public and private datasets for validating the effectiveness of the proposed method.

2. Material

Two datasets are used in this study, one is a public dataset provided by T.Posta et al. [18]., which contains 16 1-minute and 4 10-minute videos. The other is a private dataset collected from a pig farm in Foshan City, Guangdong, China, including 18 1-minute and 1 1-hour videos. Some data samples are illustrated in Fig. 1. The pigs in the public dataset are Landraces, ranging from the nursery phase (3–10 weeks old) as shown in Fig. 1(a), through the early finisher phase (11–18 weeks old) as shown in Fig. 1(b), to the late finisher phase (19–26 weeks old) as shown in Fig. 1(c). The private dataset includes Prune and Black pigs in the early finisher phase, depicted in Fig. 1(d), Fig. 1(e), and Fig. 1 (f).

To evaluate the detection and tracking performance of the algorithm under various scenes, we selected representative videos from public and private datasets to construct the test sets. The detailed data descriptions are shown in Table 1, including 20 1-minute videos, 4 10-minute videos, and 1 1-hour video with variable light during daytime and nighttime, and pig activity categorized into low, medium, and high levels denoted by L, M, and H, respectively. The videos were annotated using Darklabel software, where the annotated data contained the bounding box, ID number, and behavior categories for each pig.

3. Method

3.1. Overview of YOLOX+BoT-SORT-Slim

The main process of the proposed method is summarized in Fig. 2. In the input stage, the YOLOX detector generates detections for the video sequences, including bounding boxes and behavior categories, which are then inputted along with tracklets into the BoT-SORT-Slim module. Within the BoT-SORT-Slim module, the improved Kalman Filter(KF) initializes or predicts tracks. The tracks and detections are subsequently input into the BYTE framework for data association and tracklet management. Finally, in the output stage, the matched tracks are output, completing the behavior analysis of group-housed pigs.

3.2. YOLOX detector

YOLOX[19] is an object detector with high detection accuracy and fast inference speed. The specific architecture is illustrated in Fig. 3. The network architecture consists of three main components, the Backbone, Neck, and Head. The Backbone adopts CSPDarkNet53[20] as the feature extraction network, introduces the Path Aggregation Network (PAN) [21] in the Neck to enhance the Feature Pyramid Network (FPN), and integrates a variety of improvement techniques in the Head such as SimOTA and a Decoupled Head to create a high-performance Anchor-free detector.

In the Backbone, CBS, CSP1, CSP2, and SPP modules are employed to extract image features, resulting in three different scales of features. Subsequently, the FPN and PAN in the Neck fuse these features to obtain three multidimensional feature maps. These feature maps are then individually input into the Head, which consists of three Decoupled Detection Heads. Each detection head outputs object classes, bounding boxes, and objectness scores. Finally, the outputs from the Head undergo Concatenation and Non-Maximum Suppression to produce the final

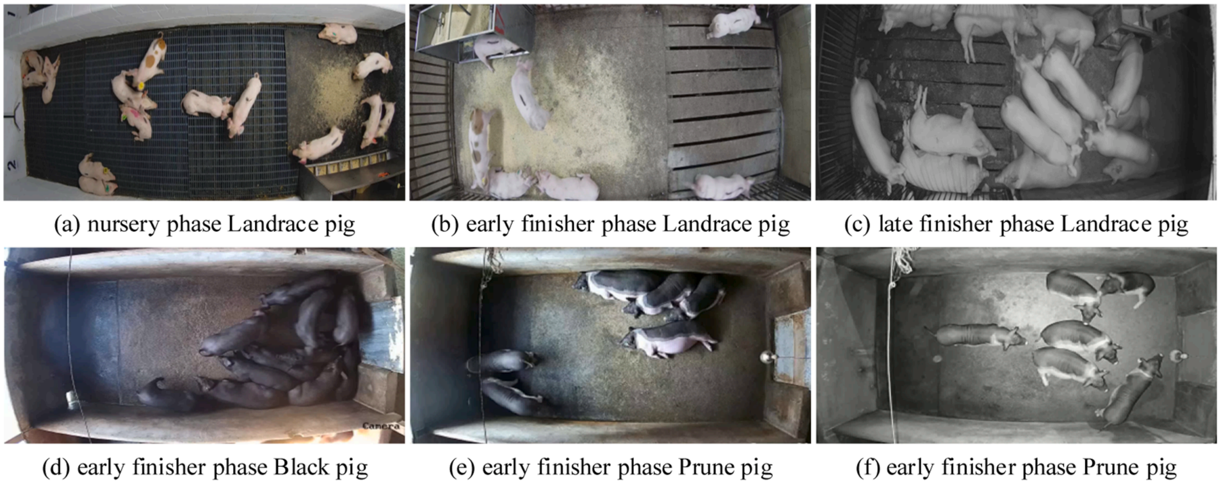


Fig. 1. Part of group-housed pig images.

Table 1
Test sets.

Public dataset	Video #	0102	0402	0502	0602	0702	0802	0902	1002	1102	1202	1502	01 10min	05 10min	11 10min	15 10min
	Day	√	√		√	√		√		√	√		√		√	
	Night			√			√		√			√		√		√
	# of Pigs	7	15	8	16	12	13	14	14	16	15	16	7	8	16	16
	Activity Level	H	L	M	M	L	L	L	L	M	L	M	M	M	H	M
Private dataset	**Video #**	**10**		**11**	**12**	**13**		**14**			**15**		**16**	**17**	**18**	**1 h**
	Day	√		√	√	√		√			√		√			√
	Night													√	√	
	# of Pigs	11		11	11	11		6			6		6	6	6	6
	Activity Level	M		M	H	H		L			L		M	M	L	H

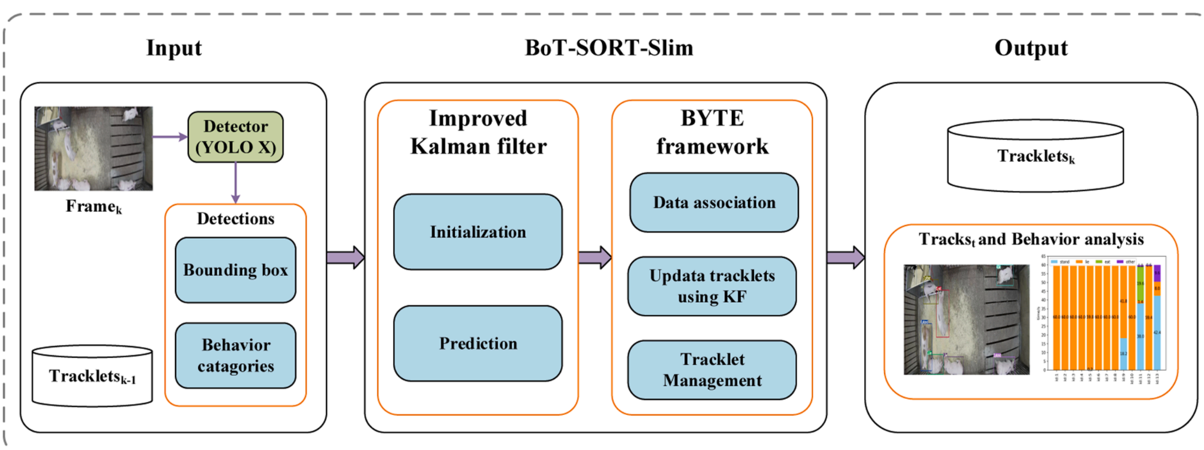


Fig. 2. The process of pig MOT and behavioral analysis.

detection results.

3.3. BoT-SORT-Slim

We propose BoT-SORT-Slim, an improved version of BoT-SORT[22]. By removing redundant operations from the original algorithm, including the motion camera compensation (CMC) and re-identification (ReID) in the group-housed pig monitoring scenario, we significantly reduce computational costs and enhance operational efficiency while maintaining tracking performance.

BoT-SORT-Slim. primarily consists of an improved KF and the BYTE framework. The KF is used to track initialization, prediction, and update. The BYTE framework is used for data association and tracklet management.

3.3.1. Improved Kalman filter

Most SORT-like algorithms adopt the KF with the constant-velocity model assumption as the motion model[22], including SORT[23] and

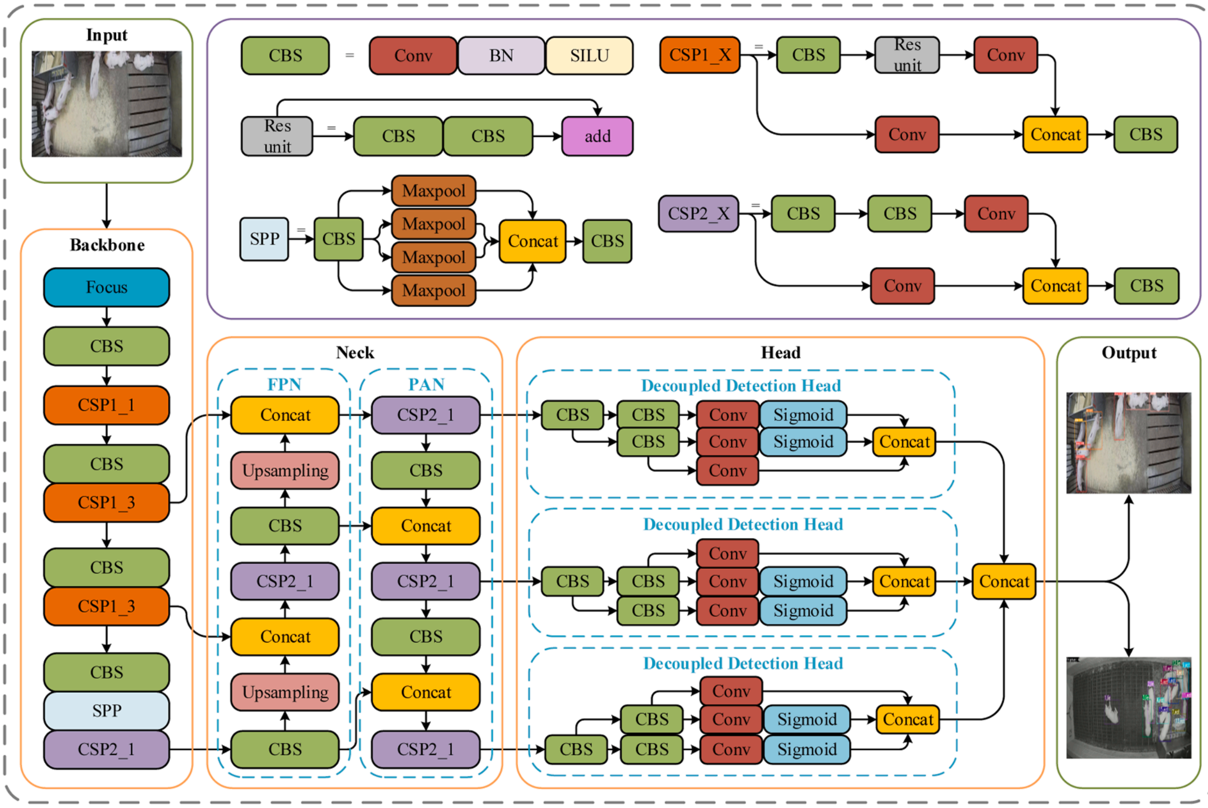


Fig. 3. The YOLOX network structure.

DeepSORT[24] et al. It has three steps in handling tracks, which are initializing, predicting, and updating trajectories. The specific steps are as follows:

(1). In the initialization phase, the bounding boxes serve as measurement vectors input into the KF and are initialized as the state vectors. In SORT the state vector was chosen to be a seven-tuple $x = [x_c, y_c, s, a, \dot{x}_c, \dot{y}_c, \dot{s}]^T$, where (x_c, y_c) are the 2D coordinates of the object center in the image plane, s is the bounding box scale (area) and a is the bounding aspect ratio. The remaining three variables represent their velocities. We expand these vectors to an 8-dimensional state vector as in Eq.(1).

$$x_k = [x_c(k), y_c(k), w(k), h(k), \dot{x}_c(k), \dot{y}_c(k), \dot{w}(k), \dot{h}(k)]^T \quad (1)$$

Where k represents the k-th frame, $(x_c(k), y_c(k))$ are the 2D coordinates of the object center in the image plane, $(w(k), h(k))$ are the width and height of the bounding box. The remaining four variables represent their velocities. The eight-dimensional state vector defined in BoT-SORT-Slim is an important improvement, enabling precise prediction of the tracking boxes, as shown in Fig. 4. The yellow boxes represent the tracking boxes using the improved KF, while the blue dashed boxes represent the tracking boxes before the improvement. It can be observed from the figure that the tracking bounding box of the improved KF can fit the target more accurately, enhancing the performance of MOT.

(2). In the prediction phase, the KF predicts the state vector x_{k-1} of pigs from the previous frame to obtain the current frame state vector $\hat{x}'_{k|k-1}$, as shown in Eq.(2). Where F_k is the KF model matrices.

$$\hat{x}'_{k|k-1} = F_k \cdot x_{k-1} \quad (2)$$

(3). In the update phase, the KF fuses the measurement vector z_k with

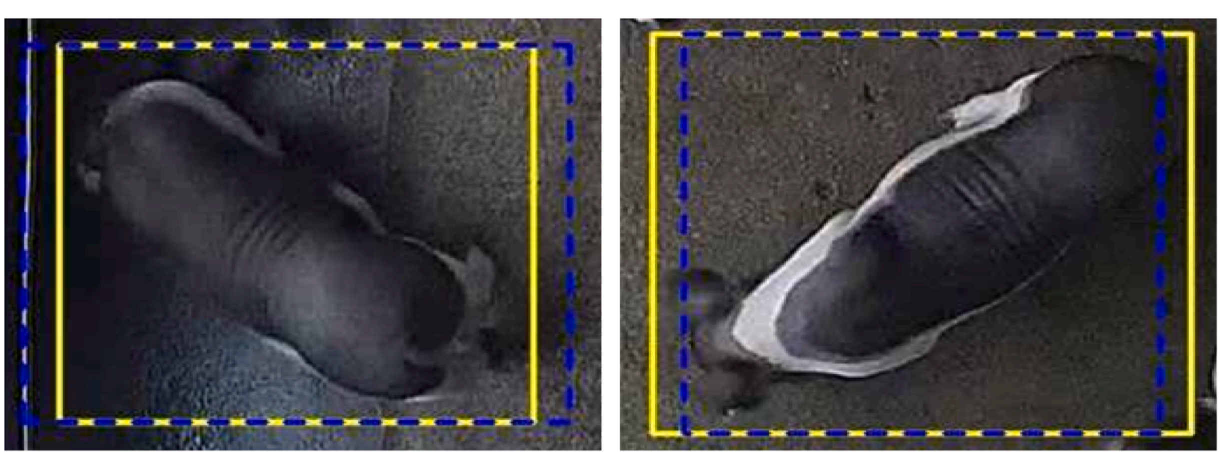


Fig. 4. The effect of the improved KF.

the state vector $\widehat{x}'_{k|k-1}$ to obtain a more accurate trajectory result, as shown in Eq.(3).

$$\widehat{x}_{k|k} = \widehat{x}'_{k|k-1} + K_k \cdot (z_k - H_k \cdot \widehat{x}'_{k|k-1}) \tag{3}$$

Where H_k is a matrix that projects the state vector $\widehat{x}'_{k|k-1}$ to the measurement space and K_k is Kalman gain, which is used to estimate the significance of the error.

3.3.2. BYTE

The BYTE is a simple, effective, and generic data association method for data association and tracklet management[25]. Its specific workflow is illustrated in Fig. 5.

(a) The input includes the detections $D_k = \{D_1, D_2, ..., D_m\}$, which be separated into two parts D_{high} and D_{low} according to the detection score threshold τ (the value in this paper is 0.65) from the current video frame and the tracklets $T_{k-1} = \{T_1, T_2, ..., T_n\}$ from the previous frame (including lost tracks T_{lost}) for two rounds of data association.
(b) The first association is performed between the high score detection boxes D_{high} and all the tracklets T_{k-1}. The similarity is computed by the IoU distances between the D_{high} and T_{k-1}. Then the unmatched tracks T_{remain} from the first association and the detections D_{low} are matched by the second IoU-based association. Finally, we get the matched tracks T_k and unmatched detection D_{remain}, tracks $T_{re-remain}$ and detections $D_{re-remain}$, and enter them all into the Tracklet management module.
(c) In the Tracklet Management module, the management of trajectories includes the following four situations. For the matched tracks T_k, update their trajectories using the KF. For the unmatched detections D_{remain}, create new tracks and assign them new ID. For the unmatched tracks $T_{re-remain}$, move them into the set T_{lost}. For the unmatched detections $D_{re-remain}$, simply delete them.

3.4. Pig behavioral analysis

We designed and implemented an analytical algorithm for statistics on the duration of four pig behaviors for overall herd behavioral analysis, the algorithm is shown in Algorithm 1. The specific implementation steps are as follows:

(1). Initialize tracks T, a pig behavior statistics dictionary A to record the number of each type of behavior for each pig as in Eq.(4), a pig behavior record dictionary B to log each pig's behavior in each frame as in Eq.(5), and a category index dictionary *category_index* as in Eq.(6).

$$A\{pig\_id : \{stand : count, lie : count, eat : count, other : count\}, \cdots\} \leftarrow \varnothing \tag{4}$$

$$B\{pig\_id : \{frame : behavior, \cdots\}, \cdots\} \leftarrow \varnothing \tag{5}$$

$$category\_index\{stand : 0, lie : 1, eat : 2, other : 3\} \tag{6}$$

For Eq.(4), if a pig's behavior is identified, add 1 to the corresponding variable. e.g., if the pig with ID 1 is identified as *stand*, add 1 to the *count* corresponding to *stand* where $pig\_id = 1$.

For Eq.(5), if the behavior of a pig in the current frame is recognized then add the corresponding record. e.g., if the behavior of the pig with ID 1 is recognized as *stand* in the first frame, add $frame = 1, behavior = stand$ to the dictionary with $pig\_id = 1$.

For Eq.(6), the corresponding index is taken out if the pig's behavior is recognized. e.g., if the behavior of the pig is identified as *stand* then $index = 1$ is obtained.

(2). Iterate through the video frames, detect the pigs, and perform data association, while assigning values to T, A, and B.

(3). Make analytical plots for the dictionaries A and B that have been processed in step (2).

4. Results and analysis

4.1. Performance and execution speed comparison of different MOT algorithms

We compared BoT-SORT-Slim with other tracking algorithms on both 1-minute public and private datasets. The tracking performance and the execution speed results are summarized in Table 2.

In terms of tracking performance, BoT-SORT-Slim achieves a HOTA of 84.7 %, MOTA of 98.3 %, IDF1 of 98.4 %, and IDs of 11 on the public dataset. In comparison, DeepSORT[24], StrongSORT[26], ByteTrack [25], OC-SORT[27], and BoT-SORT[22] have HOTA of 61.8 %, 62.3 %, 71.4 %, 82.0 %, and 84.4 %, respectively. Their MOTA are 85.1 %, 90.0 %, 90.4 %, 96.3 %, and 98.4 %, with IDF1 scores of 67.6 %, 68.0 %, 84.2 %, 96.8 %, and 98.0 %, and IDs of 327, 286, 67, 22, and 12. Among these algorithms, BoT-SORT performs the best, but our BoT-SORT-Slim shows improvement over BoT-SORT with an increase in HOTA by 0.3 % and IDF1 by 0.4 %, while reducing IDs by 11. On the private dataset, BoT-SORT-Slim achieves HOTA of 78.7 %, MOTA of 97.0 %, IDF1 of 93.4 %, and IDs of 105. Compared to the best BoT-SORT algorithm, BoT-SORT-Slim has a higher HOTA of 2.0 %, MOTA of 0.9 %, and IDF1 of 2.1 %.

In terms of execution speed, BoT-SORT-Slim achieves 68.05 FPS on the public dataset and 71.27 FPS on the private dataset, making it approximately 6.4 times and 3.6 times faster, respectively, compared to the original BoT-SORT.

The comparative tracking results between BoT-SORT-Slim and the other MOT methods on public and private datasets are presented in Fig. 6 and Fig. 7. BoT-SORT-Slim exhibits excellent tracking performance in natural scenes. As seen in Fig. 6, in frame 201 of video 0602,

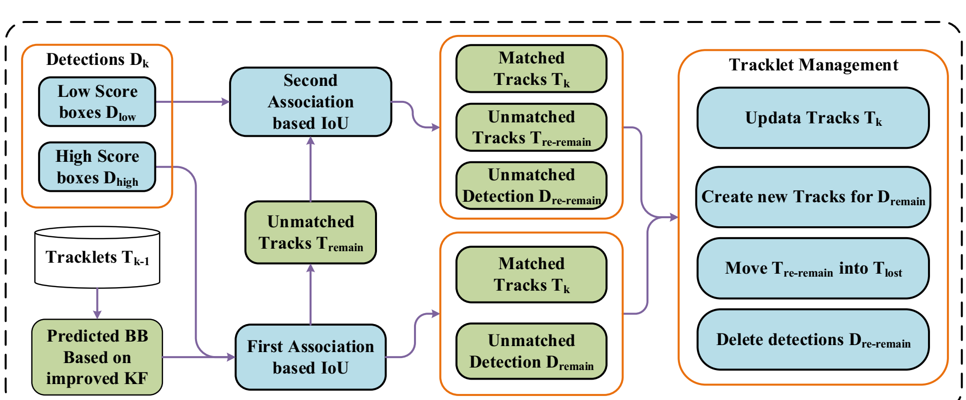


Fig. 5. Overview of BYTE framework.

Algorithm 1
Pseudo-code of Behavior Category Time Statistics of Pigs.

```
Input: A video sequence V ; object detection Det ; detection score threshold τ
Output: Tracks T of the video and dictionary A, B
Initialization: T ← ∅
                A{pig_id : {stand : count, lie : count, eat : count, other : count}, ···} ← ∅
                B{pig_id : {frame : behavior, ···}, ···} ← ∅
                category_index{stand : 0, lie : 1, eat : 2, other : 3}
for frame f_k in V do
  | D_k ← Det(f_k)
  | for d in D_k do
  |   | /* Initialize an array of behavior counts including four elements for each d */
  |   | a_d ← [0, 0, 0, 0]
  |   | /* Set a variable representing the category index*/
  |   | index ← category_index ← d_category
  |   | a_d[index] ← 1
  | end
  |
  | Associate T with D_k using BoT-SORT-Slim:
  |   | if succeed to match then
  |   |   | Call the Update or Re-activate function to update the status of tracks
  |   |   | A ← (pig_id, a_d)
  |   |   | B ← (pig_id, frame, a_d)
  |   | end
  |   | if failed to match and D.score > τ then
  |   |   | Call the function to create a new track
  |   |   | Initialize a behavior-count array a_d ← [0, 0, 0, 0]
  |   |   | A ← (pig_id, a_d)
  |   |   | B ← (pig_id, frame, a_d)
  |   | end
  | end
end
Return: T, A, and B
```

the maximum pig ID number for DeepSORT, StrongSORT, and ByteTrack reach 68, 74, and 32 respectively, which is significantly higher than 16 (the true number of pigs in the video 0602), indicating frequent error IDs. In contrast, OC-SORT and BoT-SORT-Slim have maximum ID numbers of 18 and 17 respectively. These results illustrate that BoT-SORT-Slim can obtain accurate tracking capability under natural conditions.

BoT-SORT-Slim also demonstrates superior detection and tracking performance in challenging scenarios involving uneven lighting, occlusion, and dense pig distribution. Fig. 7 shows the tracking results of video 18 for five methods, with the red dashed box indicating a missed detection and the yellow arrow indicating a false detection.

In the 1st frame where pigs are severely crowded, BoT-SORT-Slim has only 1 missed detection. In comparison, DeepSORT, StrongSORT, ByteTrack, and OC-SORT have values of missed detection and false detection as follows: 4 and 0, 2 and 0, 2 and 2, and 4 and 0 respectively. By the 11th frame, BoT-SORT-Slim successfully detects all 6 pigs without false detection, while other algorithms struggle to detect all pigs and produce false detections. By frame 120, with sparsely distributed pigs and no occlusions, OC-SORT and StrongSORT detect all pigs, whereas DeepSORT and TransTrack still fail to detect all pigs.

Considering the comprehensive analysis from Table 2, Fig. 6, and Fig. 7, BoT-SORT-Slim demonstrates superior tracking performance compared to the other four methods, making it suitable for MOT tasks in group-housed pigs.

4.2. Analysis of MOT results of BoT-SORT-Slim in each public test set

To further analyze the tracking performance of BoT-SORT-Slim, we list evaluation results for each 1-minute test set in Table 3 Tracking Results of BoT-SORT-Slim on the Public Test Set. In videos 0102, 0402, 0502, 0702, 0802, 0902, 1102, and 1202, BoT-SORT-Slim achieves HOTA of over 83 %. In videos 0602, 1002, and 1502, the HOTA are 78.8 %, 73.4 %, and 77.3 %, respectively. The lower HOTA in these videos may be attributed to higher pig activity levels and dim lighting conditions. Across all videos, MOTA consistently exceeds 96 %, with IDF1

Table 2
Comparison of performance results of Algorithms.

Datasets	Algorithms	HOTA/ %↑	MOTA/ %↑	IDF1/ %↑	IDs↓	FPS↑
Public dataset	DeepSORT	61.8	85.1	67.6	327	5.9
	StrongSORT	62.3	90.0	68.0	286	5.9
	ByteTrack	71.4	90.4	84.2	67	10.7
	OC-SORT	82.0	96.3	96.8	22	10.8
	BoT-SORT	84.4	**98.4**	98.0	12	10.5
	BoT-SORT-Slim	**84.7**	98.3	**98.4**	**11**	**68.0**
Private dataset	DeepSORT	49.0	81.8	55.4	200	9.0
	StrongSORT	57.8	85.4	67.5	379	9.6
	ByteTrack	65.0	83.7	82.5	90	17.1
	OC-SORT	74.8	93.7	93.1	**17**	17.4
	BoT-SORT	76.7	96.1	91.3	55	19.3
	BoT-SORT-Slim	**78.7**	**97.0**	**93.4**	105	**71.2**

above 97 % in most videos except for video 0102, which is 91.2 %. It also performs exceptionally well concerning IDs and execution speed, with a total of only 11 IDS across all test sets and an average FPS as high as 68.05. In conclusion, BoT-SORT-Slim ensures continuous and stable trajectories, achieving superior tracking performance.

4.3. Validation of BoT-SORT-Slim on long-term dataset

The validation results of BoT-SORT-Slim in the 10-minute and 1-hour test sets are presented in Table 4. In videos 11 and 15, the HOTA is 54.6 % and 57.9 % respectively, with high numbers of FN/FP: 1034/1561 for video 11, and 3471/645 for video 15. An analysis of these results suggests that in video 11, frequent and intense movements of the pigs led to motion blur, resulting in degraded detection and tracking performance. As in Fig. 8 video 11, pigs with ID 56 and 14 change their ID number three times within 20 frames. In Video 15, most of the pigs gather on the

Table 3
Tracking Results of BoT-SORT-Slim on the Public Test Set.

Video Sequences	HOTA/ %↑	MOTA/ %↑	IDF1/ %↑	FP↓	FN↓	IDS↓	FPS↑
0102	83.4	99.6	91.2	0	4	3	68.84
0402	89.9	99.7	99.8	0	11	0	68.06
0502	84.5	99.7	99.8	3	2	0	70.84
0602	78.8	98.6	97.4	25	41	1	63.93
0702	87.1	99.6	99.8	0	14	0	69.62
0802	92.0	99.9	99.9	0	1	0	67.90
0902	88.6	96.6	98.3	72	70	0	68.68
1002	73.4	96.6	98.2	38	103	1	68.62
1102	88.3	97.8	98.6	37	61	3	67.00
1202	85.9	97.0	98.4	56	75	2	68.28
1502	77.3	98.3	98.8	20	57	1	66.87
Total/ Average	84.7	98.3	98.4	251	439	11	68.05

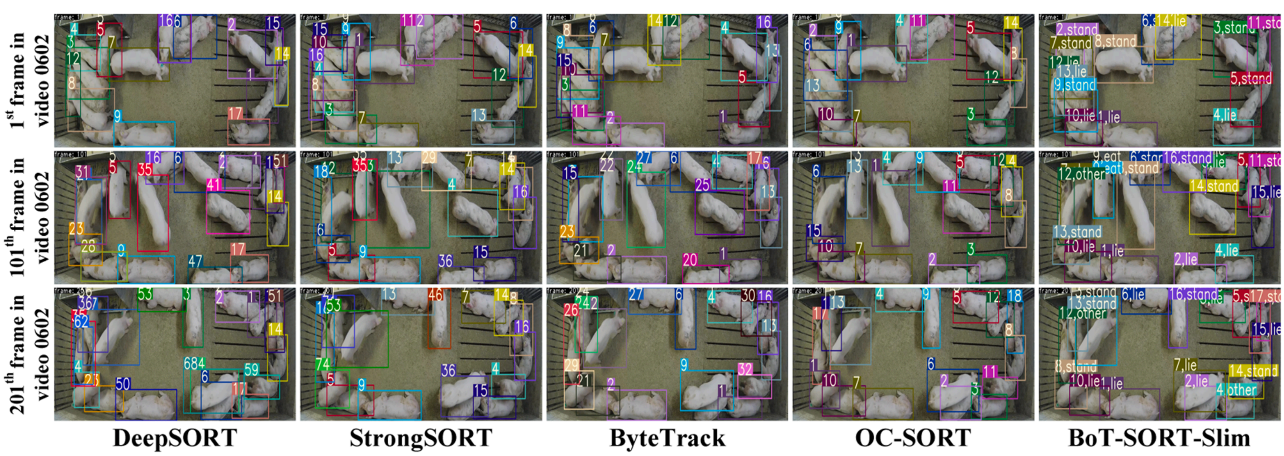


Fig. 6. Comparison of tracking performance of different methods on the public dataset.

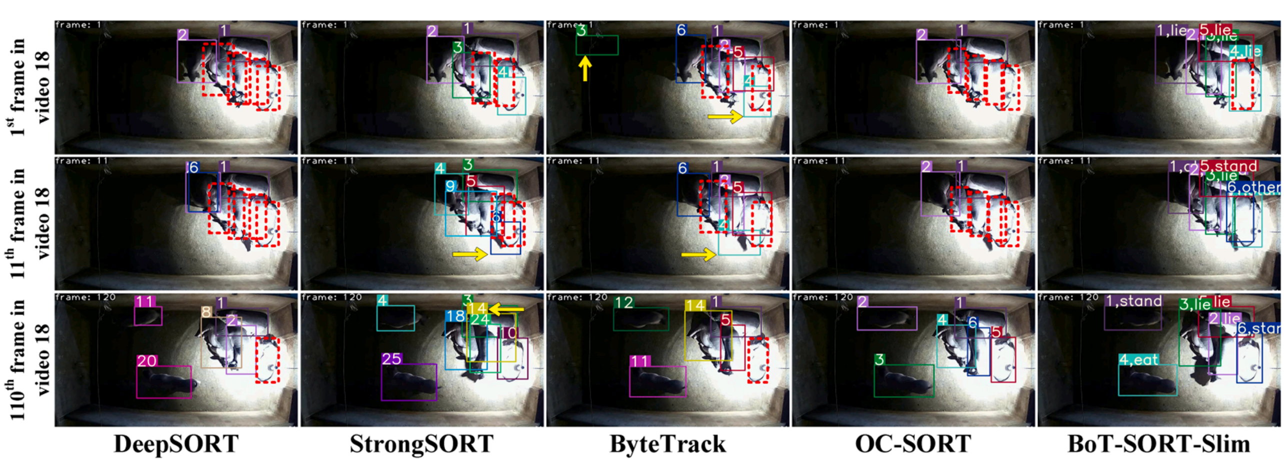


Fig. 7. Comparison of tracking performance of different methods on the private dataset.

Table 4
Validation results of BoT-SORT-Slim on the 10-minute and 1-hour Test Set.

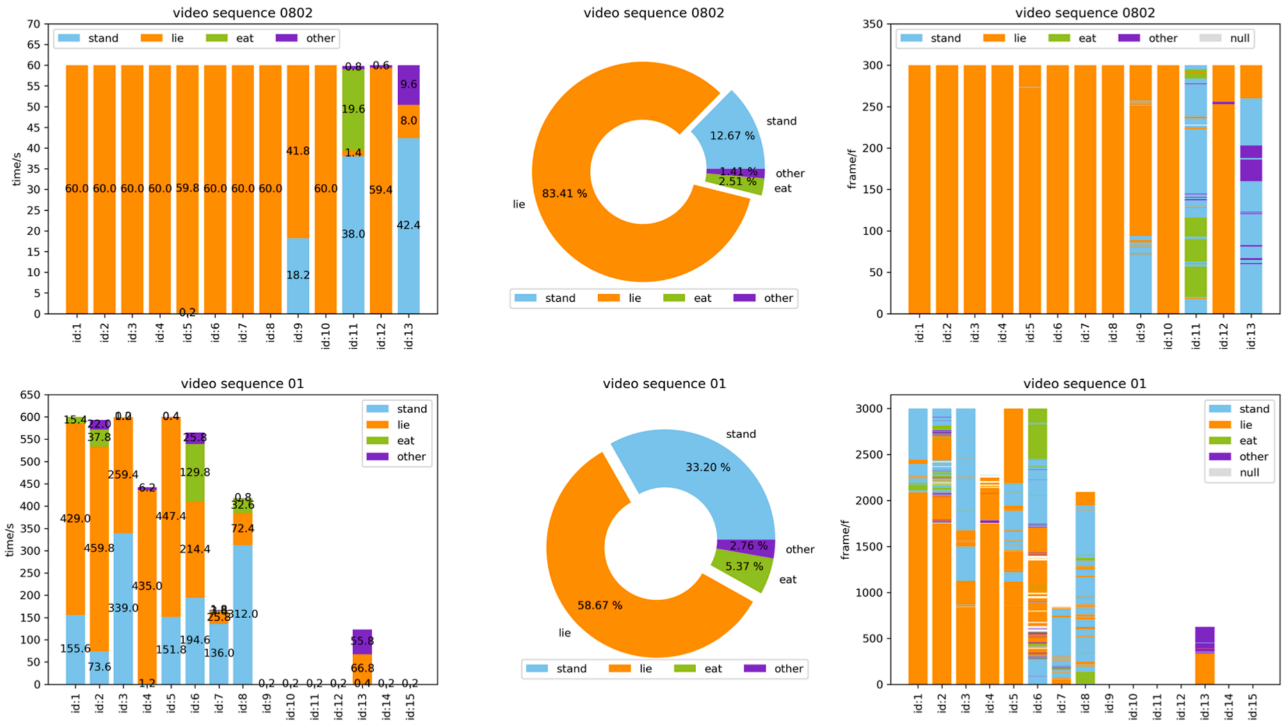
Video Sequences	HOTA/ %↑	MOTA/ %↑	IDF1/ %↑	FP↓	FN↓	IDs↓	FPS↑
01 10min	75.0	95.8	91.0	202	656	15	69.69
05 10min	77.4	98.4	93.0	32	311	19	71.60
11 10min	54.6	94.3	62.6	1561	1034	65	69.83
15 10min	57.9	91.0	63.9	645	3471	86	70.02
Total/ Average	63.3	94.1	72.5	2440	5472	185	70.28
1 h	76.4	99.6	90.2	370	53	7	72.24

right side for a long time, and the body occlusion generates error and false detections, resulting in pig IDS and trajectory loss. Therefore, achieving stable and accurate long-term tracking of pigs in complex scenarios requires further refinement of both detection and tracking algorithms.

4.4. Statistics and analysis of pig behavior

We used the proposed algorithm to analyze pig behaviors in the video to assess their health status. Some of the statistics are shown in Fig. 9, where the first row shows the behavioral statistics in the 1-minute video 0802, and the second row shows the behavioral statistics in the 10-

Fig. 8. Visualization of BoT-SORT-Slim in long-term videos.



(a) time spent by each pig in four types of behaviors **(b) percentage of collective behaviour** **(c) behavior of each pig in each frame**

Fig. 9. Pig behavioral analysis of 1-minute and 10-minute videos.

minute video 01.

Fig. 9(a) displays the time each pig spends on four different behaviors. These data can be used to improve welfare, optimize productivity, and enhance our understanding of pig behavior. For example, from Table 1, we know that video 0802 was recorded at night. Correspondingly, the figure clearly shows that most pigs are sleeping, while a few are active. For those not sleeping, we can consider their behavior to be abnormal and need to focus on their health status in the subsequent period.

Fig. 9(b) shows the percentage of all pigs involved in each type of behavior, which helps in overall behavioral assessment and individualized feeding management. For example, according to the statistics from video 0802, the "lie" behavior accounts for as much as 83.41 %. As indicated in Table 1, we can evaluate the overall activity level as low and accordingly reduce the amount of feed.

Fig. 9(c) records the behavior of each pig in every frame. By observing the detailed behavior records and videos, specific environmental factors or management measures causing abnormal pig behavior can be identified. This allows for targeted improvements in the pigs' living conditions and welfare.

Similar results were obtained for other long and short videos. In conclusion, statistical analysis of pig behavior can provide valuable information for pig farming and improve the quality of life and breeding efficiency.

5. Conclusion

This paper proposed a lightweight method for pig behavior tracking and analysis: BoT-SORT-Slim, designed for both long-term and short-term pig behavior tracking and analysis. Compared to the original YOLOX+BoT-SORT algorithm, the proposed method not only achieves superior tracking performance but also significantly improves execution speed. To support long-term behavioral tracking research and validate the robustness of our method, we manually annotated four 10-minute videos (totaling 12,000 frames) and one 1-hour video (totaling 18,000 frames).

In terms of tracking performance and execution speed, YOLOX+BoT-SORT-Slim achieved the highest HOTA (84.7 % and 78.7 %), MOTA (98.3 % and 97.0 %), IDF1 (98.4 % and 93.4 %), and FPS (68.05 and 71.27) among all tested algorithms for the 1-minute public and private test sets. The HOTA, MOTA, IDF1, and FPS in the 10-minute and 1-hour test sets were 63.3 % and 76.4 %, 94.1 % and 99.6 %, 72.5 % and 90.2 %, and 70.28 and 72.24 respectively. This indicates that our method can accurately and stably recognize and track the behavior of pigs in natural scenes.

In terms of behavioral analysis, three dimensions of statistical information were generated for each video. The behavior data can reveal the welfare status of pigs and detect abnormal behaviors. By analyzing behavior, it is possible to optimize the housing conditions and feeding regimes of group-housed pigs, thereby providing better environments and management strategies to improve production efficiency.

In conclusion, YOLOX+BoT-SORT-Slim has made significant progress in both long-term and short-term behavior tracking. The behavior analysis algorithm can also provide more intelligent and reliable solutions for the modern pig industry.

Future work

While our proposed algorithm performs well in natural and some challenging scenes, it struggles with tracking high-speed moving objects and in dimly lit conditions. Analysis of the experimental results suggests that these limitations are due to the detector's tendency to miss detections in such scenarios. In the future, we plan to enhance the detector to better address these challenges. In addition, long-time tracking (>24 h) requires more hardware and software resources, and it will be one of our future tasks to achieve more accurate pig tracking.

Data Availability

The data supporting the findings of this study are available upon request from the corresponding author.

Funding

This study was funded partly by Special Fund for the Rural Revitalization Strategy of Guangdong (2023TS-3, 2024TS-3), key R&D project of Guangzhou (202,206,010,091,2024B03J1358,2023B03J1363), Meizhou Tobacco Technology. Project of Guangdong Province (202,304).

Ethical statement

In submitting this manuscript titled "Tracking and Automatic Behavioral Analysis of Group-housed Pigs Based on BoT-SORT" to Smart Agricultural Technology, we affirm that.

Ethical approval

Any research involving human or animal subjects reported in this manuscript has been conducted in accordance with relevant institutional and national guidelines, and ethical approval has been obtained from the appropriate ethics committee(s) or institutional review board (s).

Research funding

Any sources of funding for the research reported in this manuscript have been appropriately acknowledged. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Publication ethics

We confirm that this manuscript is original, has not been published previously (except in the form of an abstract or preliminary report), and is not under consideration for publication elsewhere.

We adhere to the highest ethical standards in conducting research and preparing this manuscript for submission to Smart Agricultural Technology.

CRediT authorship contribution statement

Shuqin Tu: Writing – review & editing, Supervision, Resources, Data curation, Conceptualization. **Yuefei Cao:** Writing – original draft, Project administration, Methodology, Investigation, Formal analysis. **Yun Liang:** Investigation, Funding acquisition. **Zhixiong Zeng:** Resources. **Haoxuan Ou:** Resources. **Jiaying Du:** Resources. **Weidian Chen:** Resources.

Declaration of competing interest

All authors declare that they have no conflicts of interest that could influence the research reported in this manuscript or its interpretation.

Data availability

Data will be made available on request.

References

[1] Y. Gao, K. Yan, B. Dai, H. Sun, Y. Yin, R. Liu, W. Shen, Recognition of aggressive behavior of group-housed pigs based on CNN-GRU hybrid model with spatio-temporal attention mechanism, Comput. Electron. Agric. 205 (2023), https://doi.org/10.1016/j.compag.2022.107606.

[2] Boyle, L.A., Edwards, S.A., Bolhuis, J.E., Pol, F., Šemrov, M.Z., Schütze, S., … Valros, A.J.F.i.v.s. (2022). The evidence for a causal link between disease and damaging behavior in pigs. 8, 771682.

[3] Bowen, F., Bryant, R., & Greer, A.J.J. (2022). Behavioral Fingerprinting: acceleration Sensors for Identifying Changes in Livestock Health. 5(4), 435.

[4] Q. Guo, Y. Sun, C. Orsini, J.E. Bolhuis, J. de Vlieg, P. Bijma, P.H.N. de With, Enhanced camera-based individual pig detection and tracking for smart pig farms, in: Comput. Electron. Agric., 211, 2023, https://doi.org/10.1016/j.compag.2023.108009.

[5] W. Hao, K. Zhang, L. Zhang, M. Han, W. Hao, F. Li, G. Yang, TSML: a New Pig Behavior Recognition Method Based on Two-Stream Mutual Learning Network, Sensors. (Basel) (11) (2023) 23, https://doi.org/10.3390/s23115092.

[6] X. Yang, C. Zheng, C. Zou, H. Gan, S. Li, S. Huang, Y. Xue, A CNN-based posture change detection for lactating sow in untrimmed depth videos, Comput. Electron. Agric. 185 (2021), https://doi.org/10.1016/j.compag.2021.106139.

[7] H. Gan, J. Guo, K. Liu, X. Deng, H. Zhou, D. Luo, Y. Xue, Counting piglet suckling events using deep learning-based action density estimation, Comput. Electron. Agric. 210 (2023), https://doi.org/10.1016/j.compag.2023.107877.

[8] L. Nie, B. Li, F. Jiao, J. Shao, T. Yang, Z. Liu, ASPP-YOLOv5: a study on constructing pig facial expression recognition for heat stress, Comput. Electron. Agric. 214 (2023), https://doi.org/10.1016/j.compag.2023.108346.

[9] Y. Zhuang, K. Zhou, Z. Zhou, H. Ji, G. Teng, Systems to monitor the individual feeding and drinking behaviors of growing pigs based on machine vision, Agriculture (1) (2022) 13, https://doi.org/10.3390/agriculture13010103.

[10] Ji, H., Yu, J., Lao, F., Zhuang, Y., Wen, Y., & Teng, G.J.A. (2022). Automatic position detection and posture recognition of grouped pigs based on deep learning. 12(9), 1314.

[11] S. Tu, Y. Cai, Y. Liang, H. Lei, Y. Huang, H. Liu, D. Xiao, Tracking and monitoring of individual pig behavior based on YOLOv5-Byte, Comput. Electron. Agric. (2024) 221, https://doi.org/10.1016/j.compag.2024.108997.

[12] J. Wei, X. Tang, J. Liu, Z. Zhang, Detection of Pig Movement and Aggression Using Deep Learning Approaches, Animals. (Basel) (19) (2023) 13, https://doi.org/10.3390/ani13193074.

[13] A. Melfsen, A. Lepsien, J. Bosselmann, A. Koschmider, E. Hartung, Describing Behavior Sequences of Fattening Pigs Using Process Mining on Video Data and Automated Pig Behavior Recognition, Agriculture (8) (2023) 13, https://doi.org/10.3390/agriculture13081639.

[14] S. Tu, Q. Zeng, Y. Liang, X. Liu, L. Huang, S. Weng, Q. Huang, Automated Behavior Recognition and Tracking of Group-Housed Pigs with an Improved DeepSORT Method, Agriculture (11) (2022) 12, https://doi.org/10.3390/agriculture12111907.

[15] H. Gan, C. Xu, W. Hou, J. Guo, K. Liu, Y. Xue, Spatiotemporal graph convolutional network for automated detection and analysis of social behaviours among pre-weaning piglets, Biosyst. Eng. 217 (2022) 102–114, https://doi.org/10.1016/j.biosystemseng.2022.03.005.

[16] M. Wang, M.L.V. Larsen, D. Liu, J.F.M. Winters, J.-L. Rault, T. Norton, Towards re-identification for long-term tracking of group housed pigs, Biosyst. Eng. 222 (2022) 71–81, https://doi.org/10.1016/j.biosystemseng.2022.07.017.

[17] C. Liu, D. Wang, C. Bo, *Re*-detection and distractor association from a global perspective: a long-term tracking system, Comput. Electrical Eng. (2023) 107, https://doi.org/10.1016/j.compeleceng.2023.108611.

[18] T. Psota, E., Schmidt, T., Mote, B., & C. Pérez, L. (2020). Long-Term Tracking of Group-Housed Livestock Using Keypoint Detection and MAP Estimation for Individual Animal Identification. 20(13), 3670. Retrieved from https://www.mdpi.com/1424-8220/20/13/3670.

[19] Ge, Z., Liu, S., Wang, F., Li, Z., & Sun, J.J.a.p.a. (2021). Yolox: exceeding yolo series in 2021.

[20] Redmon, J., & Farhadi, A. (1804). YOLOv3: an Incremental Improvement. arXiv preprint arXiv:. 02767. In.

[21] S. Liu, L. Qi, H. Qin, J. Shi, J. Jia, Path aggregation network for instance segmentation, in: Paper presented at the Proceedings of the IEEE conference on computer vision and pattern recognition, 2018.

[22] Aharon, N., Orfaig, R., & Bobrovsky, B.-Z.J.a.p.a. (2022). BoT-SORT: Robust associations multi-pedestrian tracking.

[23] A. Bewley, Z. Ge, L. Ott, F. Ramos, B. Upcroft, Simple online and realtime tracking, in: Paper presented at the 2016 IEEE international conference on image processing (ICIP), 2016.

[24] N. Wojke, A. Bewley, D. Paulus, Simple online and realtime tracking with a deep association metric, in: Paper presented at the 2017 IEEE international conference on image processing (ICIP), 2017.

[25] Y. Zhang, P. Sun, Y. Jiang, D. Yu, F. Weng, Z. Yuan, X. Wang, ByteTrack: multi-object Tracking by Associating Every Detection Box, in: Computer Vision – ECCV 2022, 2022, pp. 1–21.

[26] Y. Du, Z. Zhao, Y. Song, Y. Zhao, F. Su, T. Gong, H. Meng, Strongsort: make deepsort great again, IEEe Trans. Multimedia 25 (2023) 8725–8737.

[27] J. Cao, J. Pang, X. Weng, R. Khirodkar, K. Kitani, Observation-centric sort: rethinking sort for robust multi-object tracking, in: Paper presented at the Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, 2023.