Occlusion-resistant locomotion analysis of piglets using amodal instance segmentation

H. Gan[1], A. Mao[1], C. N. Sze[1], E. Huang[1], M. C. Ceballos[2], T. D. Parsons[3] and K. Liu[1,*]
[1]*Department of Infectious Diseases and Public Health, City University of Hong Kong, Hong Kong SAR, China*
[2]*Faculty of Veterinary Medicine, University of Calgary, Calgary, Alberta, Canada*
[3]*Swine Teaching and Research Center, School of Veterinary Medicine, University of Pennsylvania, Kennett Square, PA, USA*
[*]Corresponding author: Kai Liu, kailiu@cityu.edu.hk

Abstract

Locomotion of piglets is a critical indicator of their growth, health, and welfare status; thus it is of utmost importance to automate the analysis of piglet locomotion, particularly during the early lactation periods. An intersection over unit- (IOU-) and contour-based tracking method is proposed to automate the locomotion analysis for piglets in farrowing pens. In the first step, an anchor-free deep learning network is employed in amodal instance segmentation of individual piglets. Then a novel attention graph convolution-based structure is used to distil element-wise features within the detected piglets. The distilled features are further encoded by a graph convolutional network. In the output features, pixels selected by a selection strategy derive features for a real-value pixel using 4-point nearest neighbor bilinear interpolation. Thus, a higher-resolution segmentation is predicted in a coarse-to-fine fashion. In the second step, a regional matching strategy is adopted synchronously for initial tracking. The Hungarian algorithm is then used to optimize initial tracking trajectories determined by contour features and IOUs. In the experiment, our method produced crisp amodal instance segmentation, whilst also achieving favorable tracking performance under occlusion. The tracking results demonstrated an IDF1 score of 95.7% and an MOTA of 97.1% in short video clips. With the high-accuracy occlusion-resistant segmentation and tracking results, our computer vision-based piglet tracking method may aid automated piglet locomotion and behavior analysis.

Keywords: animal welfare, occlusion, multiple object tracking, deep learning

Introduction

Piglets experiencing pain will deviate from their normal behaviors by showing abnormal locomotion which meanwhile indicates their vitality (Anil et al., 2002). Accelerometers, an electromechanical device measuring accelerating force, used to be the most widely adopted tool in analyzing pig motion and activities (Arablouei et al., 2021; Benjamin and Yik, 2019). However, piglets endure pain during the installation process of measurement devices and cannot easily carry them without damage risks (Ahrendt et al., 2011). In addition, these devices are not economically feasible for commercial piggeries that house a large number of piglets (Nasirahmadi et al., 2017). Therefore, a computer vision-based method for measuring locomotion of group-housed piglets would be more applicable and feasible.

Segmentation is considered the mainstream computer vision-based method for rough evaluation of object motion. The conventional segmentation-based motion indexes developed by Nasirahmadi et al. (2015) and Kashiha et al. (2014) were based on a fitting ellipse that required targets to be precisely segmented without occlusion. However, these activity indexes failed to represent pig motion when the body rotation was larger than 90°, as the head and hip were indistinguishable. This rough approximation of total motion is insensitive to minute-scale body movements and hampers the locomotion analysis for individual piglets.

To facilitate the locomotion analysis for individual piglets, piglets would be tracked to determine the positions between adjacent video frames. Among the deep learning-based multiple object tracking (MOT) methods, convolutional neural network (CNN) outstands for its superior performance in feature extraction (Gan et al., 2021; Sun et al., 2021). However, CNN is computationally expensive especially when it is applied in both object detection and association tasks. Intersection over unit (IOU) is a commonly used clue in pig tracking (Gan et al., 2022), yet it is not robust against complex object adhesion situation.

Due to postures and motion state, contour features differ dramatically from piglet to piglet while are close between adjacent frames for the same piglet, especially when compared within a quite limited region. To facilitate fast piglet association following piglet detection, we proposed an IOU- and contour-based tracking method for piglets in farrowing pens to automate the locomotion analysis by obtaining region differences between two adjacent frames.

Materials

Animals and video acquisition

Videos were collected at the Swine Teaching and Research Center at the University of Pennsylvania, USA. Videos from a total of seven lactating sows (Line 241; DNA Genetics, Columbus, NE) individually raised along with their litters were used in the study. Sows were constrained in a 0.64 m × 1.73 m enclosure by an open hinged farrowing stall inside a 2.1 m × 2.0 m farrowing pens during four to seven postpartum days.

An overhead camera (IPX DDK-1700D, USA) was fixed above each pen in the central position at an approximate height of 2 m, guaranteeing a bird-view of the whole pen. Videos were recorded from 7:00 am to 06:00 pm (frame rate: 7 fps; resolution: 1024×768 pixels). More information on the litters is given in Table 1. Five out of seven pens were randomly adopted as training dataset to train the amodal instance segmentation network, and the rest two pens were adopted as test dataset for the evaluation of our method. Video clips (60 s per clip) were trimmed from the 19 daytime long videos where piglets exhibited suckling, aggression, nosing, chasing, hustling, and sleeping, coupled with close contact or adhesion between piglets and occlusions from the hinged farrowing stall. Besides, an 8-h video episode from one of the test pens (i.e., Pen 3) was used for long term test. As a result, there were totally 70 60-s video clips for trainig and 30 video clips as well as one long video episode for test.

Table 1: Information on litters (Pen 3 and 5 are used for test and the rest are for training).

Pen number	1	2	3	4	5	6	7	Average
Litter size	13	11	13	9	15	8	9	11.1
Postpartum day	4, 5	4, 5, 6, 7	4, 5, 6, 7, 8	5	4, 5, 6, 7, 8	4	5	5.2

Methods

Overview of locomotion analysis of piglets

Our proposed method mainly includes two steps: 1) amodal instance segmentation to obtain contour of each piglet and 2) tracking the piglets using contour and IOU features and analyzing the locomotion (Figure 1). In the first step, an anchor-free deep learning network is employed in amodal instance segmentation for individual piglets. The localization head roughly detects regions of interest (ROI). Then a novel attention graph convolution-based structure (AGCS) is used to distil element-wise features within the detected piglets. The distilled features are further encoded by a graph convolutional network and fed into a boundary head and a mask head. The boundary head plays an auxiliary role that produces secondary supervisory signal during training phase. In the second step, the contour and IOU features are derived from instance segmentation of piglet. Then a regional matching strategy (RMS) is adopted synchronously for tracking. The Hungarian algorithm (Munkres, 1957) is then used to optimize the tracking trajectories determined by both contour and IOU features. Upon piglet tracking, locomotion analysis is conducted by using the keypoints output by a lightweighted keypoint detector.

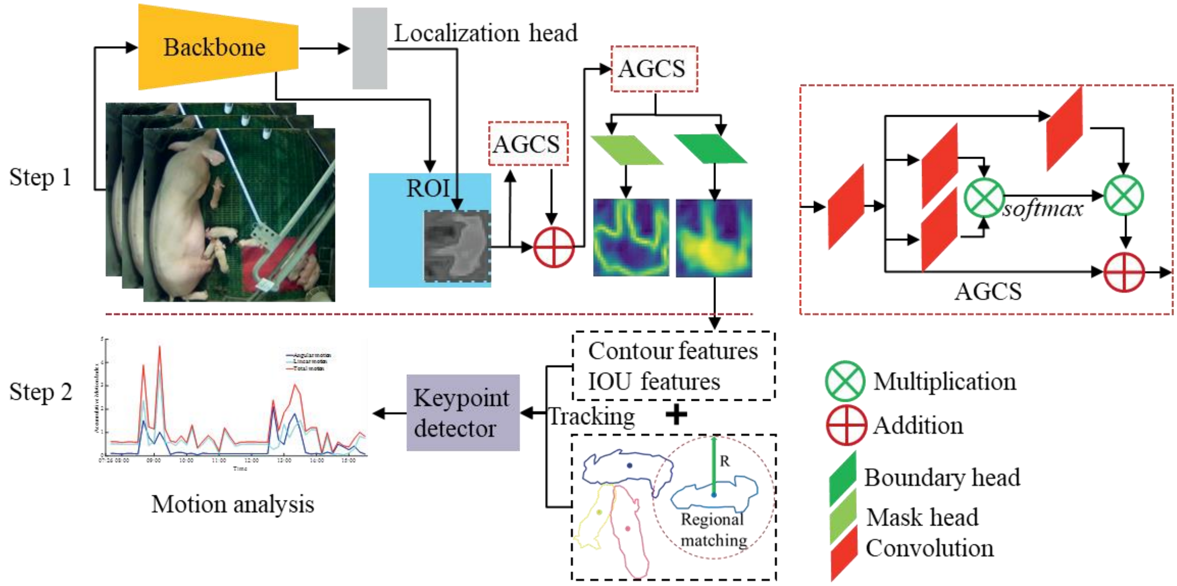

Figure 1: Workflow of the analysis of piglet locomotion.

Hu-moments for piglet instance matching

From a strictly statistical point of view, ‘moments’ are statistical expectations of a random variable. Given a silhouette of a piglet, the regular moment of a piglet instance in a binary image is defined as follows:

$$M_{!"} = \sum_{\$} \sum_{\#} x^{!} y^{"} I(x,y) \tag{1}$$

where $I(x,y)$ represents the pixel intensity value at the (x, y) coordinate. The centroid of the shape $>\tilde{x}, \tilde{y}?$is defined as follows:

$$\tilde{x} = M_{\&'}/M_{''} \tag{2}$$

$$\tilde{y} = M_{'\&}/M_{''} \tag{3}$$

Accordingly, relative moments centered about the centroid is given as follows:

$$\mu_{O} = \sum_{\$} \sum_{\#} >x - \tilde{x}? \; >y - \tilde{y}? \; I(x,y) \tag{4}$$

Hu (1962) took these relative moments and constructed seven separate moments which are suitable for shape discrimination:

$$M_{\&} = (\mu_{*'} + \mu_{'*}) \tag{5}$$

$$M_{*} = (\mu_{*'} - \mu_{'*})^{*} + 4\mu_{\&\&}^{*} \tag{6}$$

$$M_{+} = (\mu_{+'} - 3\mu_{\&*})^{*} + (3\mu_{*\&} - \mu_{+'})^{*} \tag{7}$$

$$M_{,} = (\mu_{+'} + \mu_{\&*})^{*} + (\mu_{*\&} + \mu_{'+})^{*} \tag{8}$$

$$M_{-} = (\mu_{+'} - 3\mu_{\&*})(\mu_{+'} + \mu_{\&*})((\mu_{+'} + \mu_{\&*})^{*} - 3(\mu_{*\&} + \mu_{'+})^{*}) + (3\mu_{*\&} - \mu_{'+})(\mu_{*\&} + \mu_{'+})(3(\mu_{+'} + \mu_{\&*})^{*} - (\mu_{*\&} + \mu_{'+})^{*}) \tag{9}$$

$$M_{.} = (\mu_{*'} - \mu_{'*})((\mu_{+'} + \mu_{\&*})^{*} - (\mu_{*\&} + \mu_{'+})^{*}) + 4\mu_{\&\&}(\mu_{+'} + 3\mu_{\&*})(\mu_{*\&} + \mu_{'+}) \tag{10}$$

$$M_{/} = (3\mu_{*\&} - \mu_{'+})(\mu_{+'} + \mu_{\&*})((\mu_{+'} + \mu_{\&*})^{*} - 3(\mu_{*\&} + \mu_{'+})^{*}) - (\mu_{+'} - 3\mu_{\&*})(\mu_{*\&} + \mu_{'+})(3(\mu_{+'} + \mu_{\&*})^{*} - (\mu_{*\&} + \mu_{'+})^{*}) \tag{11}$$

$$F = [M_{\&}, M_{*}, \cdots, M_{/}] \tag{12}$$

Matching score (MS) was used to compare the affinity between two piglet instances during the matching process:

$$MS = \alpha \| F_{\&} - F_{*} \| + \beta IOU \tag{13}$$

where coefficients α and β control the balance between IOU and similarity of contour features ($\| F_{\&} - F_{*} \|$). Finally, Hungarian algorithm is adopted to optimize the summing of MS values for piglets selected by RMS.

Piglet displacement between adjacent video frames is small. Accordingly, during the association of detected piglets, only those piglets having distance with current piglet lower than *ratio*L* would be targeted, where *L* is the average piglet body length in a pen and ratio is a factor controls the threshold of distance.

The motion of a piglet is composed of angular motion and linear motion. Since the snout is the most essential body part while considering social behavior, we represent the linear motion of piglets with the linear motion of the snout, instead that of the centroid. The calculation of motion index has been given in our previous study (Gan et al., 2021).

Set up

All deep learning processes were executed by using *Pytorch* framework on a single NVIDIA GeForce RTX 3090 GPU. Pretrained ResNet-101 was used as the backbone of the segmentation network. The training iteration number, initial learning rate, and the learning rate decay factor were 90k, 1×10^{-2}, and 0.1. Data was augmented by using horizontal and vertical flipping, and 180° rotation, were adopted to improve the model generalization.

As for performance indicators, $Recall_{pig}$ is adopted to evaluate the pig detection results:

$$Recall_{012} = \frac{34_{\text{pig}}}{34_{!"\#}567_{!"\#}} \tag{14}$$

where TP_{pig} is the number of detections that have IOUs > 0.5, over true bounding boxes while FP_{pig} have IOUs < 0.5, FN_{pig} denotes the number of missed piglets.

The standard COCO-style metric AP_{17} (average precision) was used to evaluate the segmentation performance (Lin et al., 2014):

$$AP_{\rho\iota} = \vartheta_{0}^{1} Precision_{\text{pixel}} Recall_{\text{pixel}} dr \tag{15}$$

where $Precision_{pixel}$ and $Recall_{pixel}$ denote the precision and recall rate of segmentation.

The MOTA and IDF_1 scores were used for piglet tracking evaluation and were calculated as follows:

$$MOTA = 1 - \frac{\sum(645675:;<)}{\sum(>3)} \tag{16}$$

$$IDF_{\&} = \frac{*:;34}{*:;345:;645:;67} \tag{17}$$

Results

Piglet instance segmentation

While using AGCS (Figure 2), performance of amodal instance segmentation is superior with a $Recall_{pig}$ of 99.0%, $AP_{0.5}$ of 97.8, and $AP_{0.75}$ of 91.7%. The occluded piglet parts were inferred accurately.

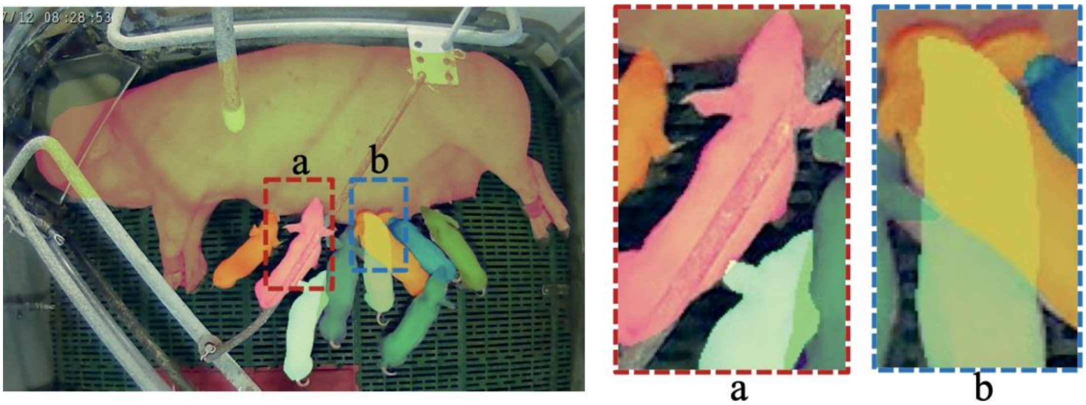

Figure 2: Amodal instance segmentation results; (a) and (b) are the enlarged views.

Piglet tracking

The optical results are obtained when α and β were 2 and 1, respectively, using a grid searching method. The tracking performance varying with searching radius is given in Figure 3. IDF1 and MOTA are highest (0.957 and 0.971) when searching ratio is 1.25.

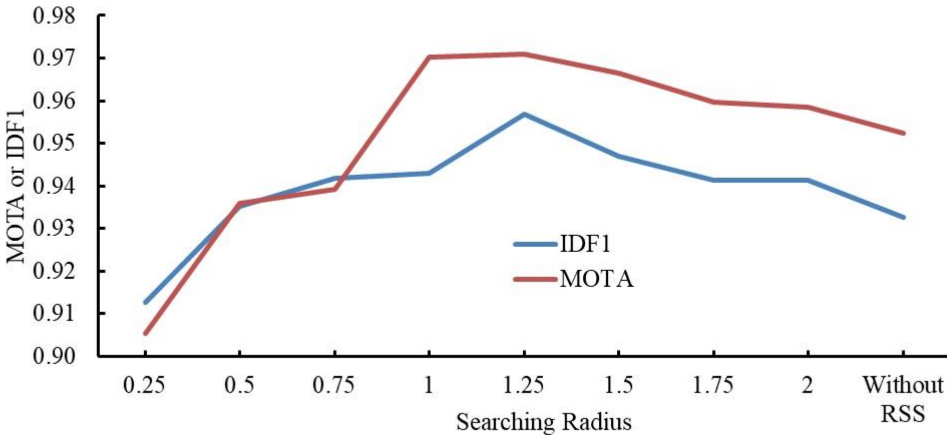

Figure 3: Tracking performance varying with searching radius.

Locomotion analysis

With the presice tracking, motion of all piglets in a pen, as well as angular motion and linear motion, is accumulated every 15 minutes which is depicted in Figure 4. It is observed that piglets are active approximately during 8:30~9:30 and 12:30~14:00 in the test video. Figure 4 shows that linear motion accounts for larger portion of the total motion than angular motion.

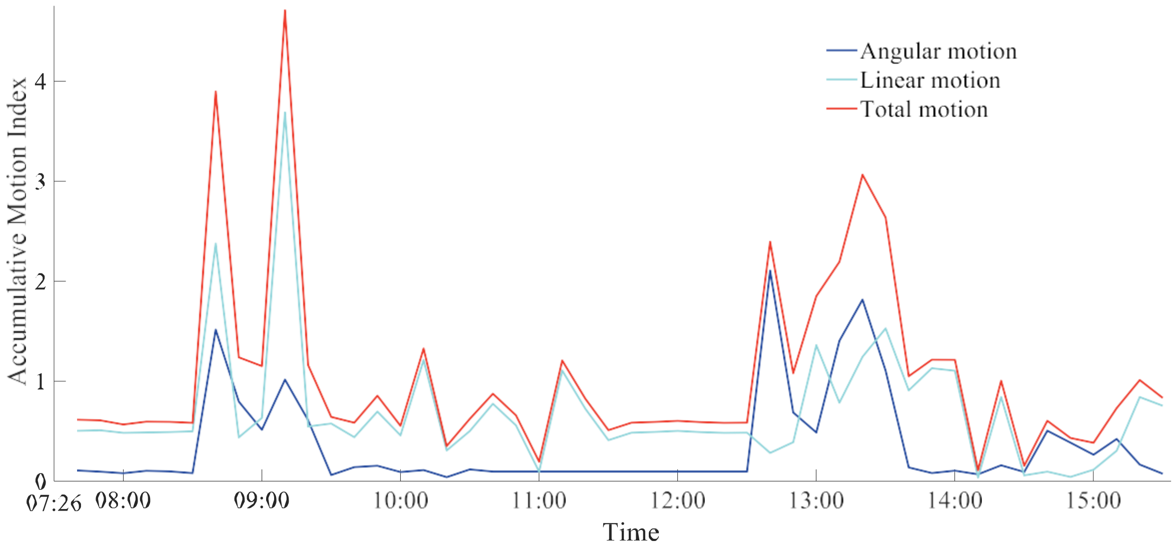

Figure 4: Piglet locomotion during the day.

Discussion

This study was conducted on small preweaning piglets with dirty and dim backgrounds and under occlusion from farrowing stalls or penmates. Piglets occluded by penmates or farrowing stalls were segmented completely and accurately when both AGCS and contour prediction were used. Unlike the conventional instance segmentation, amodal instance segmentation allows inference of the occluded part of objects, enabling our method to be occlusion-resistant.

Contour features calculated by Hu-moment are combined with IOU to facilitate fast and accurate association of piglets between adjacent video frames. Considering the small displacement of piglets between adjacent video frames, we proposed a regional searching strategy that allows piglet association within a limited region, instead of the full-scale region. It is observed when the searching radius is 1.25 times of the average piglet body length, the tracking performance is the best.

Conclusions

An intersection over unit- (IOU-) and contour-based tracking method is proposed to automate the locomotion analysis for piglets in farrowing pens. Our method produced crisp instance segmentation, whilst also achieving favorable tracking performance (an IDF1 score of 95.7% and an MOTA of 97.1%) under occlusion. With the high-accuracy occlusion-resistant segmentation and tracking results, our computer vision-based piglet tracking method may aid automated piglet locomotion and behavior analysis.

Acknowledgments

This project was funded by the new research initiatives at City University of Hong Kong (Project number: 9610450).

Animal ethical welfare statement

All animal procedures in this research study were approved by the University of Pennsylvania's Institutional Animal Care and Use Committee (Protocol #804656).

References

Ahrendt, P., Gregersen, T., and Karstoft, H. (2011) Development of a real-time computer vision system for tracking loose-housed pigs. *Computers and Electronics in Agriculture* 76(2), 169-174.

Anil, S.S., Anil, L., and Deen, J. (2002) Challenges of pain assessment in domestic animals. *Journal of the American Veterinary Medical Association* 220(3), 313-319.

Arablouei, R., Currie, L., and Kusy, B. (2021) In-situ classification of cattle behavior using accelerometry data. *Computers and Electronics in Agriculture* 183, 106045.

Benjamin, M., and Yik, S. (2019) Precision Livestock Farming in Swine Welfare: A Review for Swine Practitioners. *Animals* 9(4), Article 4.

Gan, H., Ou, M., Huang, E., Xu, C., Li, S., Li, J., Liu, K., and Xue, Y. (2021) Automated detection and analysis of social behaviors among preweaning piglets using key point-based spatial and temporal features.
Computers and Electronics in Agriculture 188, 106357.

Gan, H., Ou, M., Li, C., Wang, X., Guo, J., Mao, A., Ceballos, M.C., Parsons, T.D., Liu, K., and Xue, Y. (2022) Automated detection and analysis of piglet suckling behaviour using high-accuracy amodal instance segmentation. *Computers and Electronics in Agriculture* 199, 107162.

Gan, H., Ou, M., Zhao, F., Xu, C., Li, S., Chen, C., and Xue, Y. (2021) Automated piglet tracking using a single convolutional neural network. *Biosystems Engineering* 205, 48-63.

Hripcsak, G., and Rothschild, A.S. (2005) Agreement, the F-Measure, and Reliability in Information Retrieval. *Journal of the American Medical Informatics Association* 12(3), 296-298.

Hu, M.K. (1962) Visual pattern recognition by moment invariants. *IRE Transactions on Information Theory* 8(2), 179-187.

Kashiha, M.A., Bahr, C., Ott, S., Moons, C.P.H., Niewold, T.A., Tuyttens, F., and Berckmans, D. (2014) Automatic monitoring of pig locomotion using image analysis. *Livestock Science* 159, 141-148.

Lee, J., Jin, L., Park, D., and Chung, Y. (2016) Automatic Recognition of Aggressive Behavior in Pigs Using a Kinect Depth Sensor. *Sensors* 16(5), Article 5.

Lin, T.-Y., Maire, M., Belongie, S., Hays, J., Perona, P., Ramanan, D., Dollár, P., and Zitnick, C.L (2014) Microsoft COCO: Common Objects in Context. *In Computer Vision–ECCV 2014: 13th European Conference* Zurich, Switzerland, 740-755.

Munkres, J. (1957) Algorithms for the Assignment and Transportation Problems. *Journal of the Society for Industrial and Applied Mathematics* 5(1), 32-38.

Nasirahmadi, A., Edwards, S.A., and Sturm, B. (2017) Implementation of machine vision for detecting behaviour of cattle and pigs. *Livestock Science* 202, 25-38.

Nasirahmadi, A., Richter, U., Hensel, O., Edwards, S., and Sturm, B. (2015) Using machine vision for investigation of changes in pig group lying patterns. *Computers and Electronics in Agriculture* 119, 184-190.

Sun, S., Akhtar, N., Song, H., Mian, A., and Shah, M. (2021) Deep Affinity Network for Multiple Object Tracking. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 43(1), 104-119.