



Review

Behaviour recognition of pigs and cattle: Journey from computer vision to deep learning

Chen Chen^{a,b,*}, Weixing Zhu^{a,*}, Tomas Norton^{b,*}^a School of Electrical and Information Engineering, Jiangsu University, Zhenjiang 212013, Jiangsu, China^b Division of Measure, Model & Manage Bioresponses (M3-Biores), KU Leuven, Kasteelpark Arenberg 30, 3001 Leuven, Belgium

ARTICLE INFO

Keywords:

Identification
 Livestock behaviour recognition
 Computer vision
 Deep learning
 Research trend

ABSTRACT

The increasing demand for sustainable livestock products also demands new considerations in animal breeding. Breeding programs are now seeking to integrate animal behavioural phenotypes, as these relate to the productivity, health and welfare of the animals and thereby can influence yield and economic benefits in the industry. Traditional manual observation of pig behaviour is time-consuming, laborious, subjective, and difficult to achieve in continuous and large-scale operations. It is not surprising that computer vision technology with the advantages of being objective, non-invasive and continuous has been widely researched for its use in the recognition of livestock behaviours over recent years. Nevertheless, in studies of livestock behaviour recognition, computer vision technology faces some challenges, e.g., complex scenes, variable illumination, occlusion, touching and overlapping between livestock, which has limited the fast translation of technology to industry. On the other hand, deep learning technology has proven to solve these difficulties to a certain extent and is being adopted to recognise livestock behaviours. This paper mainly evaluates the recent developments in computer vision methods for recognition of these behaviours in pigs and cattle. The focus on these species is made possible by the number of studies exist quantifying behaviours that are of importance for their health, welfare and productivity such as aggression, drinking, feeding, lameness, mounting, posture, tail-biting and nursing. This review paper especially analyses the development of image segmentation, identification and behaviour recognition using tradition computer vision and more recent deep learning methods, and evaluates the evolution of key research in the field. We elaborate the research trend of livestock behaviour recognition from four aspects, i. e., development of robust livestock identification algorithms, recognition of livestock behaviours for different growth stages, further quantification of the results of behaviour recognition, and building evaluation system of growth status, health and welfare.

1. Introduction

With the increasing demand for sustainable animal products, livestock breeding and careful animal management have become an important means of improving production efficiency of the livestock industry (Norton et al., 2019). Livestock behaviours reflect the health, welfare and growth status of the pigs, thereby affecting the yield and economic benefits (Larsen et al., 2021). Individual animal behaviours are related to the amount of water and feed consumed, and are important for understanding animal productivity (Botreau et al., 2007). The social behaviour of animals can give important insights into their

welfare status, for example aggression between pigs can cause skin trauma, infection and even fatal injuries (Turner et al., 2006). Excessive mounting behaviours can cause a high risk of poor welfare, arising from skin lesions, lameness and stress, and economic losses from reduced performance (Teixeira and Boyle, 2014). Tail-biting is considered to be a welfare-reducing problem with economic consequences for pig production (Larsen et al., 2019). Playing behaviour of pigs towards enrichment objects can reduce the occurrence of tail-biting, mounting and aggression and consequently improve animal welfare (Lahrmann et al., 2018). Nursing behavior, as one of behaviours of sows during lactation, is critical for early survival and growth of their piglets prior to

* Corresponding authors at: School of Electrical and Information Engineering, Jiangsu University, Zhenjiang 212013, Jiangsu, China (C. Chen). School of Electrical and Information Engineering, Jiangsu University, Zhenjiang 212013, Jiangsu, China (W. Zhu). Division of Measure, Model & Manage Bioresponses (M3-Biores), KU Leuven, Kasteelpark Arenberg 30, 3001 Leuven, Belgium (T. Norton).

E-mail addresses: ainicc1987@163.com (C. Chen), wzxhu@ujs.edu.cn (W. Zhu), tomas.norton@kuleuven.be (T. Norton).

<https://doi.org/10.1016/j.compag.2021.106255>

Received 26 January 2021; Accepted 7 June 2021

Available online 15 June 2021

0168-1699/© 2021 Elsevier B.V. All rights reserved.

weaning (Muns et al., 2013), which has a great impact on the economic benefit of pig farms (Vila and Tummaruk, 2016). Furthermore, animal body-part movement can be used for disease detection, for example lameness has become a frequent and serious problem for herd productivity and animal welfare in the dairy industry (Bruijnjs et al., 2012). Clinical lameness has a significant impact on milk yield (Ouaed et al., 2015) and reproductive performance (Morris et al., 2011). Cattle pose estimation is a key step analysing cattle behaviors and evaluating cattle health, thereby greatly significant for intelligent breeding of cattle (Yazdanbakhsh et al., 2017). Therefore, the monitoring and recognition of livestock behaviours are of great significance to the development of precision livestock farming.

As the traditional method of manual observation of livestock behaviours is time-consuming, laborious, subjective, and difficult to achieve continuous and large-scale operations, in recent years computer vision technology that has the advantages of objective, non-invasive and continuous has been widely applied for recognition of livestock behaviours, e.g., aggression, drinking, feeding, lameness, mounting, posture, tail-biting, nursing, playing and other behaviours. Computer vision is a simulation of biological vision using computers and related equipment, and is an important part of the field of artificial intelligence. Its main task is to obtain the information of the corresponding scene by processing the collected images or videos. A traditional computer vision system is mainly aimed at extracting features from images, and it also includes a series of other subtasks, e.g., edge detection, corner detection, image segmentation and pattern recognition. Traditional feature extraction algorithms include Scale Invariant Feature Transform (SIFT), Speeded Up Robust Features (SURF), and Binary Robust Independent Element Feature (BRIEF) (García et al., 2020). According to the type and quality of input images, different algorithms have different degrees of success. Ultimately, the accuracy of the entire system depends on the method of extracting features. The main challenge of this method is to tell the system which features to look for in the image. As the algorithm runs according to the designer's definition and the extracted features are artificially designed, poor algorithm performance can be improved by fine-tuning in the implementation. However, such changes need to be done manually and hard-coded for specific applications, which pose a big obstacle to the realisation of high-quality computer vision. In the existing studies of livestock behaviour recognition, computer vision technology also faces other challenges, e.g., complex scenes, variable illumination, occlusion, touching and overlapping between livestock (Tu et al., 2020).

Deep learning technology that can solve the above difficulties to a certain extent has been gradually used to recognise livestock behaviours. For instances, Yang et al. (2018a) used a fully convolutional network (FCN) to segment images of lactating sows with different scenes, variable illumination, etc. Furthermore, Tian et al. (2019) used the modified Counting Convolutional Neural Network (CNN) model based on the architecture ResNeXt to count the number of pigs under the conditions of partial occlusion, overlapping and different perspectives. Therefore, development from computer vision to deep learning is necessary in the field of livestock behaviour recognition. Currently, deep learning systems have made significant progress in dealing with some related subtasks. The biggest difference in deep learning is that it no longer uses carefully programmed algorithms to search for specific features, but instead trains the neural network in the deep learning system (Yang & Xiao, 2020). As the computing power provided by deep learning systems increases, the computer will be able to recognise and react to everything it sees, which has made significant progress. In recent years, the development of deep learning has not only break-through many difficult visual problems to improve the level of image cognition, but also accelerated the progress of related technologies in the field of computer vision. With the continuous improvement of deep learning models and computing power, autonomous systems can continue to develop steadily and truly realise the interpretation and response to what they perceive.

Through the investigation of a large number of computer vision-

based and deep learning-based pig and cattle behaviour recognition literature, this paper mainly evaluates the methods for recognition of these behaviours of pigs and cattle. Since image segmentation and identification are the basis of livestock behaviour recognition, this paper analyses the development process of image segmentation, identification and behaviour recognition from computer vision to deep learning, and provides researchers and producers with technical references. Furthermore, this paper elaborates the research trend of livestock behaviour recognition from four aspects, i.e., development of robust livestock identification algorithms, recognition of livestock behaviours for different growth stages, further quantification of the results of behaviour recognition, and building evaluation system of growth status, health and welfare.

This paper is organised as follows: Section 2 reviews the studies of image segmentation of pigs and cattle from body entirety to body part. Section 3 reviews the studies of identification of pigs and cattle from body part to body entirety. Section 4 reviews the studies of behaviour recognition of pigs and cattle from computer vision to deep learning. Section 5 proposes the research trend of livestock behaviour recognition.

2. Image segmentation

The role of image segmentation is to extract foreground targets from the background, and the effect of image segmentation directly affects the accuracy of feature extraction and livestock behavior recognition (Hao et al., 2020). Table 1 illustrates the overview of image segmentation of pigs and cattle based on computer vision and deep learning.

2.1. Body entirety segmentation

Initially, the segmentation of pigs and cattle was focused on the body entirety. On the one hand, the body entirety segmentation mainly involves single-pig segmentation. In order to segment the sow and its piglets in farrowing pens (Fig. 1(a)), Tu et al. (2013) updated the background modelling as the reference image by using texture integration and then performed subtraction between the current image and the current reference image obtained by a wavelet transform. Finally, the pairwise relationships between a pixel and its neighbours on a factor graph were modeled based on the pseudo-wavelet coefficients, and the image probabilities were approximated by using loopy belief propagation. This study showed how it is possible to extract foreground pigs in complex farrowing pen scenes, e.g., sudden light changes, dynamic background, and motionless foreground objects. In order to segment the single-sow in farrowing crates during day and night (Fig. 1(b)), Khoramshahi et al. (2014) defined a number of image local features based on small neighborhoods around a point and compared their separability and performance metrics. Finally, a feed-forward neural network (NN) was used to classify the pixels between the sow and background, and a realistic configuration in terms of an acceptable level of accuracy and computation time was chosen. This study demonstrated the possibility of foreground extraction when there is lack of information of the background. The above methods belong to computer vision. In order to further separate the single-sow from the pig herd in the loose pen (Fig. 1(c)), Yang et al. (2018a) segmented the sow by using a fully convolutional network (FCN) and refined the coarse output of the FCN by using the probability map from the final layer of the FCN and Otsu's thresholding from the hue, saturation, and value colour information. This deep learning method improved segmentation results compared with Simultaneous Detection and Segmentation (SDS), Otsu, Mixture of Gaussians (MoG), and traditional FCNs. In the above studies, it can be seen that single-sow entirety segmentation has evolved from computer vision to deep learning.

On the other hand, the body entirety segmentation mainly involves multi-pig/cattle segmentation. In order to extract pigs in group-housed environments (Fig. 1(d)), Guo et al. (2014) detected the foreground pigs by using Prediction Mechanism-based Mixture of Gaussians (PM-MoG)

Table 1
Overview of image segmentation of pigs and cattle based on computer vision and deep learning.

Year	Persons	Species	Research contents	Perspectives/Cameras	Technologies
2013	Tu et al.	pig	Body entirety segmentation	Top-view/2D	Computer vision
2014	Khoramshahi et al.	pig	Body entirety segmentation	Top-view/2D	Computer vision
2014	Guo et al.	pig	Body entirety segmentation	Top-view/2D	Computer vision
2015	Guo et al.	pig	Body entirety segmentation	Top-view/2D	Computer vision
2016	Lu et al.	pig	Body entirety segmentation	Top-view/2D	Computer vision
2018	Brünger et al.	pig	Body entirety segmentation	Top-view/2D	Computer vision
2019a	Li et al.	pig	Body entirety segmentation	Top-view/2D	Computer vision
2018a	Yang et al.	pig	Body entirety segmentation	Top-view/2D	Deep learning
2019a	Zhang et al.	pig	Body entirety segmentation	Top-view/2D	Deep learning
2020	Tu et al.	pig	Body entirety segmentation	Top&side-view/2D	Deep learning
2019a	Qiao et al.	Cattle	Body entirety segmentation	Top&side-view/2D	Deep learning
2020	Xu et al.	Cattle	Body entirety segmentation	Top&side-view/2D	Deep learning
2021	Jia et al.	Cattle	Body part segmentation	Side&back-view/3D	Computer vision
2019	Psota et al.	pig	Body part segmentation	Top-view/2D	Deep learning
2019	Jiang et al.	Cattle	Body part segmentation	Side-view/2D	Deep learning
2020a	Liu et al.	Cattle	Body part segmentation	Side-view/2D	Deep learning

and used the maximum entropy segmentation according to the colour information of foreground objects. Finally, the results of the two previous steps of foreground detection were fused. This method can extract relatively complete foreground pigs in complex scenes, e.g., light changes, the influence of ground urine stains, water stains, manure and other sundries, pigs' slow movement patterns, and varying colours of pigs. In order to further extract pigs in a drinker and feeder zone (Fig. 1(e)), Guo et al. (2015) segmented pigs with a maximum entropy global threshold from the image enhanced by histogram equalization. Then, each object centroid was calculated from the initially segmented objects, and the original image was adaptively divided into multiple circular sub-blocks whose origin was the centroid and radius was the maximum distance from the centroid to the edge point. Finally, an accurate secondary segmentation result was obtained using multilevel thresholding segmentation in each sub-block. However, the above two methods will be useless when the pigs are touching each other. In order to separate these touching pigs, Lu et al. (2016) developed a segmentation algorithm for touching pig images based on ellipse fitting method (Fig. 1(f)). First, ellipse fitting was used for a large number of images which have one pig. Parameters range of ellipses fitted by images with a single piglet on different age was extracted. Secondly, contours of connected components in a touching pigs' image were extracted. Each contour was segmented based on concave points. Ellipse fitting was implemented for each contour segment. Finally, 5 rules for ellipse merging were proposed, which were used to merge anomalous ellipses. After ellipse merging, the number of ellipses was equal to the number of pigs. However, this algorithm's accuracy may be reduced when the number of pigs is more than 7. In order to further separate touching pigs, Li et al. (2019a) developed a pig detection approach based on multi-feature template matching (Fig. 1(g)). First, two complementary descriptors, i. e., dominant orientation templates (DOTs) and brightness ratio templates (BRTs), were combined to represent images. A DOT established relationships with edge features, while a BRT encoded intensity and texture features. Both descriptors were given in binary form to increase the time efficiency of the algorithm. Then, representative pig templates can be obtained from training images through an automatic selection process based on template clustering. Finally, the input image was scanned to compute matching responses for all templates. Thus, objects having similar visual properties as the given templates were considered as detected pigs. This study can effectively and reliably extract pigs in complex scenes including multiple sources of disturbances, e.g., uneven illumination, foreground objects of varying colour, pigs moving slowly, etc. The above methods of multi-pig segmentation belong to computer vision. As the deep learning method can segment pigs with good performance in the conditions of pigs' touching, overlapping and occlusion, different perspectives, etc., many studies of the body entirety segmentation of multi-pig/cattle based on deep learning have appeared.

Brünger et al. (2018) introduced a method for adapting the ellipses of pigs, which is not based on the edges of the segmentation but looks at all segmented pixels (Fig. 1(h)). This makes it easier to compensate minor errors in segmentation and helps to process images even under sub-optimal conditions, such as poor lighting or unfavourable camera positioning. Zhang et al. (2019a) segment pigs by using the Single Shot Multibox Detector (SSD) architecture with a modified loss function as the detector and compared the Faster Region-based Convolutional Neural Network (Faster R-CNN), Region-based Fully Convolutional Network (R-FCN) and SSD for pig detection (Fig. 1(i)). This study can detect pigs in conditions of light fluctuation, similar appearances of pigs, shape deformations, and occlusions and can track pigs under both daylight and infrared (nighttime) light conditions. Tu et al. (2020) developed a segmentation approach based on Mask Scoring R-CNN framework to solve pig segmentation from front-view and top-view (Fig. 1(j)). First, a CNN-based feature extraction, which used backbone architecture with residual network of depth 101 layers (ResNet-101) and feature pyramid network (FPN) that can extract features from low-level and high-level of the feature pyramid according to different scales. Then, the candidate regions of interest (ROIs) were generated by region proposal network (RPN) based on the features derived from backbone architecture. Finally, the pig for each ROI was detected and segmented. This approach can robustly detect and segment multiple pigs under the group-housed pig natural scenes, e.g., uneven illumination, pigs' occlusion and overlapping. For body entirety segmentation of multi-cattle, Qiao et al. (2019a) developed a segmentation approach based on a Mask R-CNN deep learning framework to solve problems of cattle segmentation and contour extraction in real feedlot environment (Fig. 1(k)). This approach consisted of 4 steps, i.e., key frame extraction used for detecting the huge cattle motion frames, image enhancement used for reducing the illumination and shadow influence, cattle segmentation, and body contour extraction. This study can render fairly desirable cattle segmentation performance with 0.92 Mean Pixel Accuracy (MPA) and achieve contour extraction with an Average Distance Error (ADE) of 33.56 pixels, which was better than that of the state-of-the-art Sharp-Mask and DeepMask segmentation methods. Xu et al. (2020) applied the cutting-edge instance segmentation framework of Mask R-CNN to segment quadcopter acquired cattle images and to count cattle in different situations such as extensive production pastures and also in intensive housing such as feedlots (Fig. 1(l)). The optimal IoU threshold and the full-appearance detection for the algorithm in this study were verified through performance evaluation. This study showed the framework's potential to perform reliably in offline quadcopter vision systems in counting cattle on pastures and in feedlots. Compared with the existing typical competing algorithms, Mask R-CNN outperformed both in the counting accuracy and average precision especially on the datasets with occlusion and overlapping. In the above studies, it can be

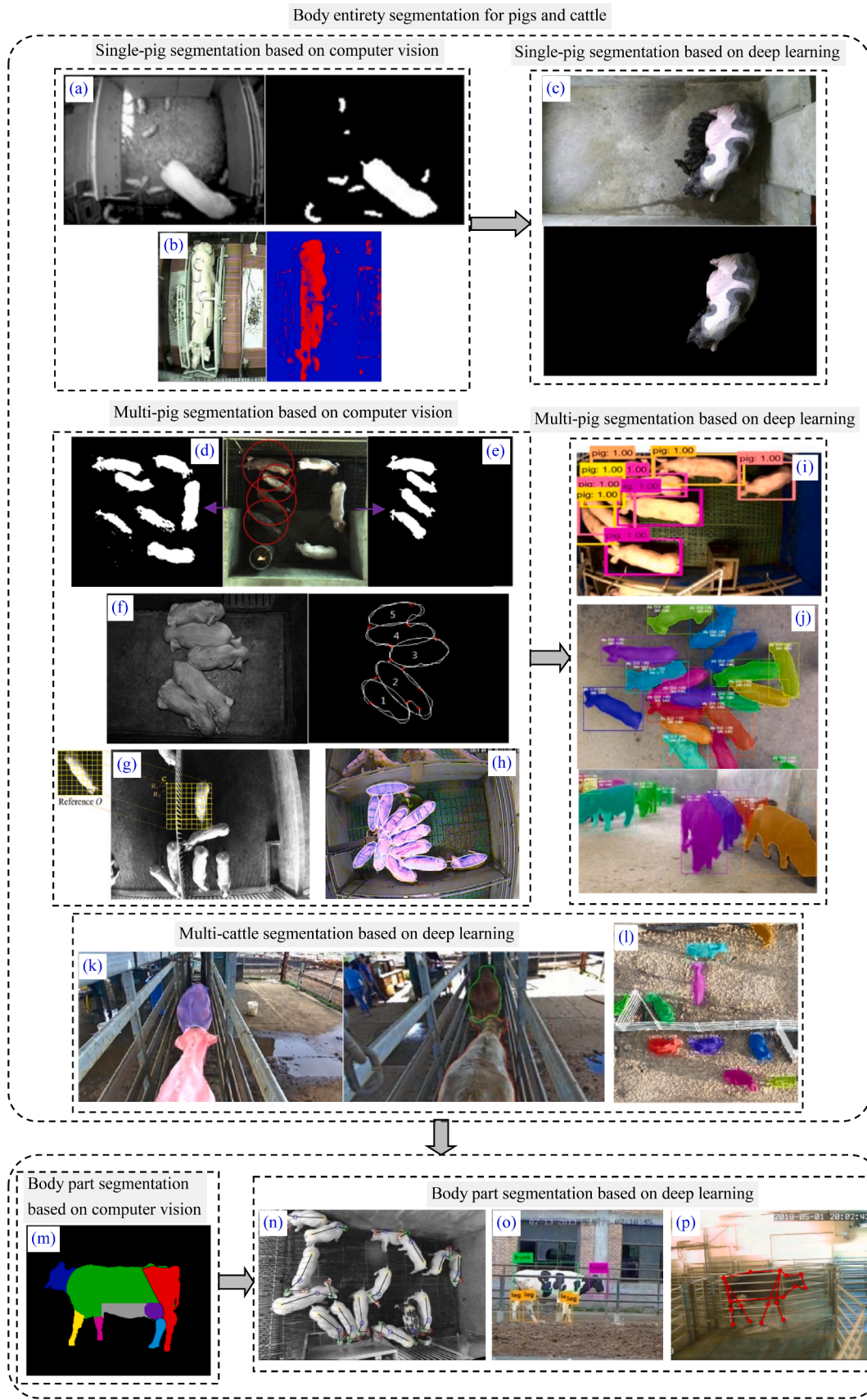


Fig. 1. Image segmentation for pigs and cattle based on computer vision and deep learning.

seen that multi-pig/cattle entirety segmentation has also developed from computer vision to deep learning.

2.2. Body part segmentation

Subsequently, the segmentation of pigs and cattle was focused on the body parts, which can provide detailed information for behaviour recognition. Jia et al. (2021) presented a computer vision method that automatically segments nine cow's body parts, i.e. head, torso, udder, belly (or rear), left foreleg, right foreleg, left hindleg, right hindleg and tail, by using an improved version of template matching (Fig. 1(m)). The results indicate that it is possible to automatically detect and extract body parts from RGB-depth images without any manual interference. Psota et al. (2019) developed an algorithm for instance-level detection of multiple pigs in group-housed environments (Fig. 1(n)). This method used a single fully-convolutional neural network to detect the location of left ear, right ear, shoulder and tail and the orientation of each pig, where both body part locations and pairwise associations were represented in the image space. In order to overcome the difficulty that traditional object detection algorithms cannot detect the key parts of dairy cows with high precision, Jiang et al. (2019) tested a deep learning network named FLYOLOv3 based on FilterLayer (FilterLayer YOLOv3) to detect dairy cows' key parts (i.e., head, trunk and legs) in complex scenes (Fig. 1(o)). As images were unstable during the training process and initialisation, particle noise was generated in feature maps after convolution. Therefore, the mean filtering algorithm was carried out, and a leaky rectifier function (Leaky ReLU) was used to integrate the custom FilterLayer layer to reduce training interference. The artificial annotation was used firstly to mark the borders of the cow's head, trunk and legs in images, and then the FLYOLOv3 network was trained with the labelled samples. Finally, the trained model was used to test the images. This method was compared with the Faster R-CNN and YOLOv3 algorithm by using indicators, e.g., accuracy, recall rate, average frame rate, and average accuracy. Liu et al. (2020a) designed a side-view cow structural model to describe the spatial positions of the joints (key-points) of the cow and developed a system using deep learning to automatically extract the structural model from videos (Fig. 1(p)). This system can detect multiple cows in the same frame and provides robust performance for the body region under practical challenges like obstacles (fences) and poor illumination. In the above studies, it can be seen that multi-pig/cattle part segmentation mainly adopts deep learning technologies.

3. Identification

The role of identification is to determine the identity of each individual in the group, and this identity can locate the recognised behaviour on a specific animal, thereby realising the transformation from group behaviour recognition to individual behaviour recognition (Prashanth & Sudarshan, 2020). Table 2 illustrates the overview of identification of pigs and cattle based on computer vision and deep learning.

3.1. Pig identification from body part to entirety

Pig identification from body part to entirety has developed from computer vision to deep learning. Ahrendt et al. (2011) developed a real-time computer vision system for identification and tracking of pigs in loose-housed stables (Fig. 2(a)). The tracking algorithm operates in 2 steps. The first step built up support maps, pointing to preliminary pig segments in each video frame. In the second step the support map segments were used to build up a 5D-Gaussian model of the individual pigs (i.e. position and shape). The system has software correction for fisheye distortion coming from the camera lens. The fisheye lens allows the camera to monitor a much larger area in the stable. The system can track at least 3 pigs more than 8 min without losing track and identity of the individual pigs in a realistic experiment. However, this study requires

Table 2

Overview of identification of pigs and cattle based on computer vision and deep learning.

Year	Persons	Species	Research contents	Perspectives/ Cameras	Technologies
2011	Ahrendt et al.	pig	Body part-based identification	Top-view/2D	Computer vision
2018	Hansen et al.	pig	Body part-based identification	Top&side-view/2D	Deep learning
2020	Marsot et al.	pig	Body part-based identification	Top&side-view/2D	Deep learning
2017	Zhu et al.	pig	Body entirety-based identification	Top-view/2D	Computer vision
2017	Huang et al.	pig	Body entirety-based identification	Top-view/2D	Computer vision
2020	Huang et al.	pig	Body entirety-based identification	Top-view/2D	Computer vision
2018b	Yang et al.	pig	Body entirety-based identification	Top-view/2D	Deep learning
2008	Allen et al.	Cattle	Body part-based identification	Top&side-view/2D	Computer vision
2014	Lu et al.	Cattle	Body part-based identification	Top&side-view/2D	Computer vision
2016	Gaber et al.	Cattle	Body part-based identification	Top&side-view/2D	Computer vision
2017	Li et al.	Cattle	Body part-based identification	Top-view/2D	Computer vision
2018	Kumar et al.	Cattle	Body part-based identification	Top&side-view/2D	Deep learning
2019	Zhao et al.	Cattle	Body part-based identification	Side-view/2D	Computer vision
2019	Sun et al.	Cattle	Body part-based identification	Top-view/2D	Deep learning
2020	Hu et al.	Cattle	Body part-based identification	Top&side-view/2D	Deep learning
2019	Okura et al.	Cattle	Body entirety-based identification	Top&side-view/2D/3D	Computer vision
2019b	Qiao et al.	Cattle	Body entirety-based identification	Top&side-view/2D	Deep learning

the manual marking on the region of pig back. In order to realise the pig identification without manual marking, Hansen et al. (2018) captured the pig face images in a farm environment and trained a CNN model using an artificially augmented data set (Fig. 2(b)). Class Activation Mapping using Grad-CAM was used to show the regions that the network used to discriminate between pigs. This study further compared the proposed CNN model with the Fisherfaces and the VGG-Face pre-trained face CNN model, which are techniques adopted in human face recognition. Marsot et al. (2020) further detected pig faces and eyes by two Haar feature-based cascade classifiers and one shallow convolutional neural network to extra high-quality images (Fig. 2(c)). Face recognition was performed by using a deep convolutional neural network, and class activation maps generated by grad-CAM and saliency maps were utilised to visually understand how the discriminating parameters have been learned by the neural network. The outcome of this study will facilitate the real-application of AI-based animal identification in swine production. However, these pig face-based identification methods require that

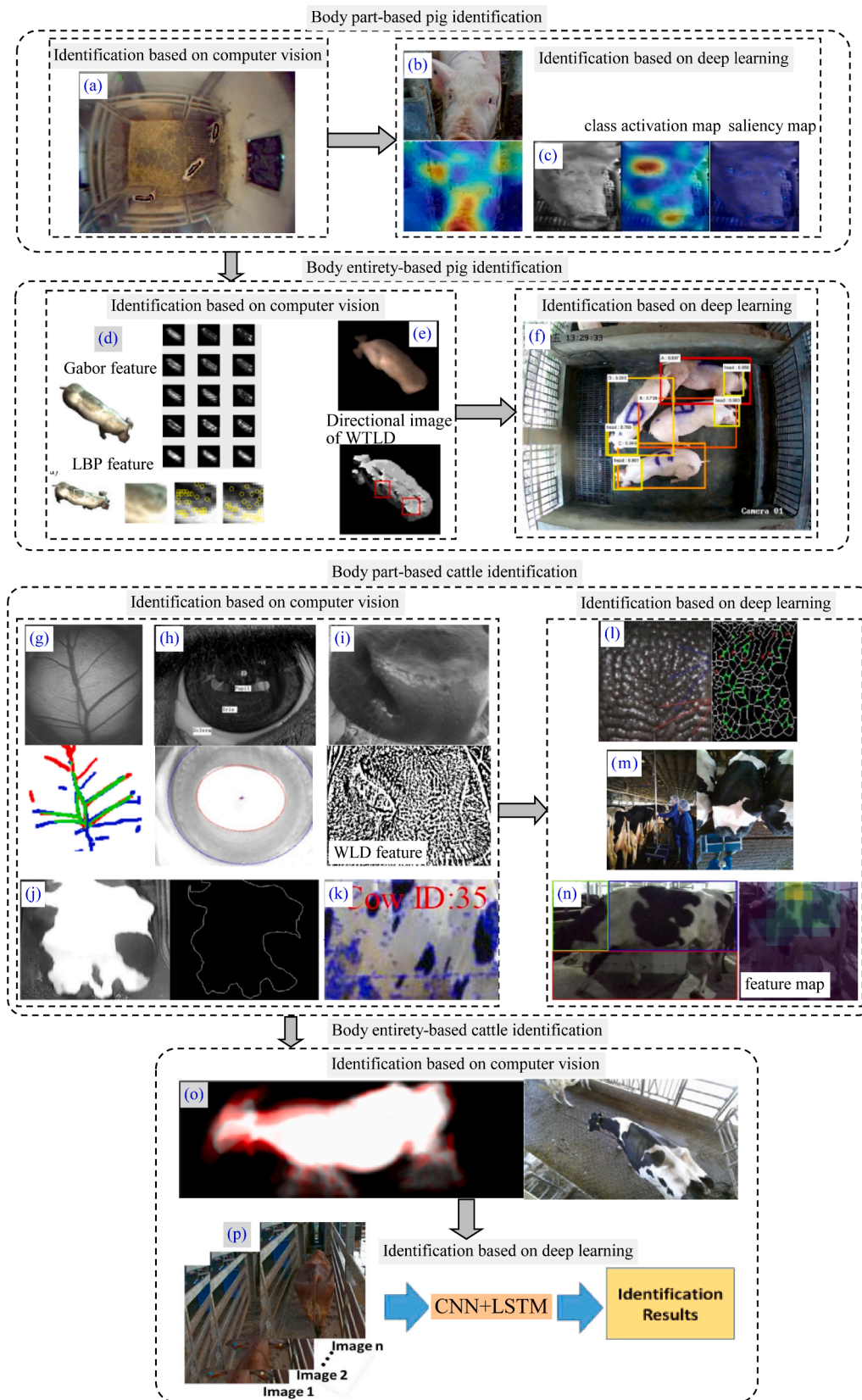


Fig. 2. Identification for pigs and cattle based on computer vision and deep learning.

pigs enter the designated area cooperatively in order to obtain images of pig faces, which may involve human intervention.

In order to realise pig identification without human intervention, the region on pigs used for identification has moved from the body part to the body entirety. [Zhu et al. \(2017\)](#) extracted the colour moments, area, perimeter and other features of the drinking pig and identified individual pigs by computing the Euclidean distance between the pig and the standard sample. [Huang et al. \(2018\)](#) extracted Gabor features by convolving pig images with Gabor filters and the local structural features by using the Local Binary Pattern (LBP) ([Fig. 2\(d\)](#)). Principal Components Analysis (PCA) was then used to reduce the feature dimension and the features were concatenated to form the feature vectors. Finally, Support Vector Machine (SVM) was used to classify these feature vectors to identify standing pigs. [Huang et al. \(2020\)](#) proposed a Weber texture local descriptor (WTLD) for the identification of pigs by extracting the local features of back hair, skin texture and spots ([Fig. 2\(e\)](#)). By calculating the differential excitation and multi-directional information of pixels, the local structure features of the main direction were fused to enhance the description ability of features. The results show that the WTLD achieves higher recognition rates with a lower feature dimension. This method can identify pig individuals with different positions and postures in the pigsty. However, the above computer vision methods for identification of pigs are mainly suitable for loose pens. In order to further identify touching pigs, [Yang et al. \(2018b\)](#) used Faster R-CNN to locate and identify individual pigs from a group-housed pen ([Fig. 2\(f\)](#)). The head of each pig was also located. An algorithm for associating the head of each pig with its body was designed. In the above studies, it can be seen that pig identification based on either body part or entirety has developed from computer vision to deep learning.

3.2. Cattle identification from body part to entirety

Cattle identification from body part to entirety has also developed from computer vision to deep learning. In order to identify the cattle based on body part, [Allen et al. \(2008\)](#) recorded the unique retinal vascular pattern (RVP) at the back of a cattle's eye and matched the RVP with the same eye of the same animal by the Optibrand Matching Engine ([Fig. 2\(g\)](#)). [Lu et al. \(2014\)](#) developed a cow identification system based on iris analysis, which includes iris imaging, iris detection, and recognition ([Fig. 2\(h\)](#)). First, the image quality of the captured sequences was assessed and a clear iris image was selected for subsequent process. Second, the inner and outer boundaries of cow iris were fitted respectively as two ellipses based on the edge images during segmentation. Then the segmented cow iris was obtained and normalised by using geometric method. Finally, 2D complex wavelet transform (2D-CWT) was used to extract local and global characteristics of the cow iris and the phase of the filtered cow iris was encoded as features for matching the identity of the cow. [Gaber et al. \(2016\)](#) used the Weber Local Descriptor (WLD) to extract robust features from cattle muzzle print images ([Fig. 2\(i\)](#)). Then the AdaBoost classifier was used to identify head of cattle from their WLD features. This WLD along with AdaBoost algorithm gave promising results compared to both of the k -Nearest Neighbor (k -NN) and Fuzzy- k -Nearest Neighbor (Fk-NN) algorithms. In the above studies, body part used for identification was mainly focused on the cattle's retinal, eye or nose. This requires these body parts being in constrained environments. In order to identify the cattle in unconstrained environments, [Li et al. \(2017\)](#) used the tailhead image as a Region of Interest (ROI) and then used Zernike moments as descriptors of shape characteristics for the white pattern on the ROI ([Fig. 2\(j\)](#)). Two groups of Zernike moments were extracted from the preprocessed image and classified using four alternative classifiers, namely, linear discriminant analysis (LDA), quadratic discriminant analysis (QDA), artificial neural network (ANN) and SVM. These results show that the low-order Zernike moment feature, along with the QDA and SVM algorithms is an effective approach for individual dairy cow identification and has

significant applications in precision animal management. [Zhao et al. \(2019\)](#) developed a computer vision system to extract body images and identify Holstein cows in the side-view videos of dairy cattle walking in a straight line ([Fig. 2\(k\)](#)). Cow mask was detected using adaptive SOM method. The largest inscribed rectangle was extracted to locate the cow's body area. Feature points of the body image were extracted and matched with the template dataset to identify unknown cows. Four feature extraction methods and two matching methods were investigated and evaluated. The results showed that the highest identification accuracy generated when the Features from Accelerated Segment Test (FAST), Scale Invariant Feature Transformation (SIFT) and FLANN methods were used for feature extraction, descriptor and matching. However, the combination of Oriented FAST and Rotated BRIEF (ORB) and BruteForce had better computational efficiency on the basis of high accuracy. The above studies belong to computer vision methods for cattle identification based on body part. On the other hand, deep learning has also been used for cattle identification based on body part. [Kumar et al. \(2018\)](#) developed a deep learning-based approach for identification of individual cattle based on their muzzle point image pattern ([Fig. 2\(l\)](#)). The deep learning-based Convolutional Neural Network (CNN) and Deep Belief Network (DBN) approaches were used to extract the salient set of texture features and represent muzzle point image of cattle. Stacked Denoising Auto-encoders (SDAE) was used to encode the extracted feature of muzzle point images. This deep learning approach outperforms state-of-the-art methods for identification of cattle on muzzle point image database. [Sun et al. \(2019\)](#) developed an automatic system for identifying individual cattle and assessing body condition score (BCS) using a deep learning framework ([Fig. 2\(m\)](#)). This work developed a linear regression model of BCS using ultrasound backfat thickness to determine BCS for training sets and tested a system based on convolutional neural networks with 3 channels, including depth, gray, and phase congruency. [Hu et al. \(2020\)](#) developed a cow identification method based on fusion of deep parts features ([Fig. 2\(n\)](#)). First, a set of side-view images of cows were captured, and then the YOLO object detection model was applied to locate the cow object in each original image, which was then divided into three parts, i.e., head, trunk and legs, by a part segmentation algorithm using frame differencing and segmentation span analysis. Then, three independent convolutional neural networks (CNNs) were trained to extract deep features from these three parts, and a feature fusion strategy was designed to fuse the features, i.e., deep parts feature fusion. Finally, a support vector machine (SVM) classifier trained by the fused features was used to identify each individual cow. In the above studies, it can be seen that the body part-based cattle identification has developed from computer vision to deep learning.

Furthermore, studies of the cattle identification based on body entirety have appeared. [Okura et al. \(2019\)](#) developed a method for cow identification based on three-dimensional video analysis using RGB-D cameras, which capture images with RGB colour information and subject distance from the camera. Using RGB-D videos of walking cows, a unified approach using two complementary features for identification, gait (i.e., walking style) and texture (i.e., markings), was developed ([Fig. 2\(o\)](#)). [Qiao et al. \(2019b\)](#) developed a deep learning based framework to identify beef cattle using image sequences unifying the advantages of both CNN (Convolutional Neural Network) and LSTM (Long Short-Term Memory) network methods ([Fig. 2\(p\)](#)). A CNN network (i.e., InceptionV3) was used to extract features from a rear-view cattle video dataset and these extracted features were then used to train an LSTM model to capture temporal information and identify each individual animal. In the above studies, it can be seen that the body entirety-based cattle identification has also developed from computer vision to deep learning.

4. Behaviour recognition

[Table 3](#) illustrates the overview of behaviour recognition of pigs and

Table 3
Overview of behaviour recognition of pigs and cattle based on computer vision and deep learning.

Year	Persons	Species	Research contents	Perspectives/Cameras	Technologies
2014a	Viazzi et al.	pig	Aggression recognition	Top-view/2D	Computer vision
2014	Oczak et al.	pig	Aggression recognition	Top-view/2D	Computer vision
2016	Lee et al.	pig	Aggression recognition	Top-view/3D	Computer vision
2017	Chen et al.	pig	Aggression recognition	Top-view/2D	Computer vision
2018	Chen et al.	pig	Aggression recognition	Top-view/2D	Computer vision
2019	Chen et al.	pig	Aggression recognition	Top-view/3D	Computer vision
2020a	Chen et al.	pig	Aggression recognition	Top-view/2D	Deep learning
2013	Kashiha et al.	pig	Drinking recognition	Top-view/2D	Computer vision
2016	Lao et al.	pig	Drinking recognition	Top-view/3D	Computer vision
2017	Zhu et al.	pig	Drinking recognition	Top-view/2D	Computer vision
2019	Leonard et al.	pig	Drinking recognition	Top-view/3D	Computer vision
2019b	Zhang et al.	pig	Drinking recognition	Top-view/2D	Deep learning
2020	Yang et al.	pig	Drinking recognition	Top-view/2D	Deep learning
2020b	Chen et al.	pig	Drinking recognition	Top-view/2D	Deep learning
2016	Lao et al.	pig	Feeding recognition	Top-view/3D	Computer vision
2018b	Yang et al.	pig	Feeding recognition	Top-view/2D	Deep learning
2019	Leonard et al.	pig	Feeding recognition	Top-view/3D	Computer vision
2020	Yang et al.	pig	Feeding recognition	Top-view/2D	Deep learning
2020	Alameer et al.	pig	Feeding recognition	Top-view/2D	Deep learning
2020c	Chen et al.	pig	Feeding recognition	Top-view/2D	Deep learning
2015	Porto et al.	Cattle	Feeding recognition	Top-view/2D	Computer vision
2020	Achour et al.	Cattle	Feeding recognition	Top-view/2D	Deep learning
2020	Bezen et al.	Cattle	Feeding recognition	Top-view/3D	Deep learning
2008	Song et al.	Cattle	Lameness recognition	Side-view/2D	Computer vision
2010	Poursaberi et al.	Cattle	Lameness recognition	Side-view/2D	Computer vision
2013	Viazzi et al.	Cattle	Lameness recognition	Side view/2D	Computer vision
2014b	Viazzi et al.	Cattle	Lameness recognition	Top&side-view/2D/3D	Computer vision
2014	Herterm et al.	Cattle	Lameness recognition	Top-view/3D	Computer vision
2018	Zhao et al.	Cattle	Lameness recognition	Side -view/2D	Computer vision
2020	Wu et al.	Cattle	Lameness recognition	Side -view/2D	Deep learning
2020	Jiang et al.	Cattle	Lameness recognition	Side -view/2D	Deep learning
2016	Nasirahmadi et al.	pig	Mounting recognition	Top-view/2D	Computer vision
2019	Guo et al.	Cattle	Mounting recognition	Top&side-view/2D	Computer vision
2019b	Zhang et al.	Pig	Mounting recognition	Top-view/2D	Deep learning
2019b	Li et al.	Pig	Mounting recognition	Top-view/2D	Deep learning
2013	Porto	Cattle	Behaviour posture recognition	Top-view/2D	Computer vision
2015	Nasirahmadi et al.	pig	Behaviour posture recognition	Top-view/2D	Computer vision
2019a	Nasirahmadi et al.	pig	Behaviour posture recognition	Top-view/2D	Computer vision
2018	Zheng et al.	pig	Behaviour posture recognition	Top-view/3D	Deep learning
2019b	Nasirahmadi et al.	pig	Behaviour posture recognition	Top-view/2D	Deep learning
2019c	Li et al.	Cattle	Behaviour posture recognition	Side -view/2D	Deep learning
2020	Riekert et al.	pig	Behaviour posture recognition	Top&side-view/2D	Deep learning
2020	Zhu et al.	pig	Behaviour posture recognition	Top-view/3D	Deep learning
2020	Zheng et al.	pig	Behaviour posture recognition	Top-view/3D	Deep learning
2020b	Liu et al.	pig	Tail-biting recognition	Top-view/2D	Deep learning
2018c	Yang et al.	pig	Nursing recognition	Top-view/2D	Deep learning
2019	Yang et al.	pig	Nursing recognition	Top-view/2D	Deep learning
2020	Yang et al.	pig	Nursing recognition	Top-view/2D	Deep learning
2020d	Chen et al.	Pig	Playing recognition	Top-view/2D	Deep learning
2016	Lao et al.	pig	Multi-behaviour recognition	Top-view/3D	Computer vision
2019	Leonard et al.	pig	Multi-behaviour recognition	Top-view/3D	Computer vision
2019b	Zhang et al.	Pig	Multi-behaviour recognition	Top-view/2D	Deep learning
2020	Yang et al.	pig	Multi-behaviour recognition	Top-view/2D	Deep learning
2020	Zhang et al.	pig	Multi-behaviour recognition	Top&side-view/2D	Deep learning
2020	Li et al.	pig	Multi-behaviour recognition	Top&side-view/2D	Deep learning
2020	Fuentes et al.	Cattle	Multi-behaviour recognition	Top&side-view/3D	Deep learning
2020	Yin et al.	Cattle	Multi-behaviour recognition	Top&side-view/2D	Deep learning

cattle based on computer vision and deep learning.

4.1. Aggressive behaviour recognition

Initially, [Viazzi et al. \(2014a\)](#) used the Motion History Image (MHI) to extract the number of moving pixels of the pig herd as the mean intensity of motion and extract the ratio of the moving pixels to the herd's pixels as the occupation index ([Fig. 3\(a\)](#)). Based on these two features, the Linear Discriminant Analysis (LDA) was used to recognise aggressive behaviour of pigs. [Oczak et al. \(2014\)](#) extracted the activity index of the pig herd and calculated the average, maximum, minimum, sum and variance of activity index as features. A multilayer feed forward neural network was trained and validated to classify events of high aggression and medium aggression. [Lee et al. \(2016\)](#) used a Kinect depth sensor to

extract the minimum, maximum, average, standard deviation of velocity, and distance between the standing pigs as feature vectors ([Fig. 3\(b\)](#)). A two binary-classifier SVM in a hierarchical manner was used to classify these feature vectors to recognise pig aggression. However, the pig herd or the standing pigs include the moving but non-aggressive pigs, and thus the features extracted from these pigs will influence the accuracy of aggression recognition. In order to further separate the pairwise aggressive pigs form the pig herd, [Chen et al. \(2017\)](#) tracked 2 aggressive pigs by using connected area and adhesion index according to their characteristics of continuous and large-proportion adhesion in the aggression process and then considered these 2 pigs as an entire rectangle to extract the acceleration feature of this rectangle ([Fig. 3\(c\)](#)). Based on this feature, rules for aggression recognition were designed to recognise medium and high aggression. In order to further extract actual

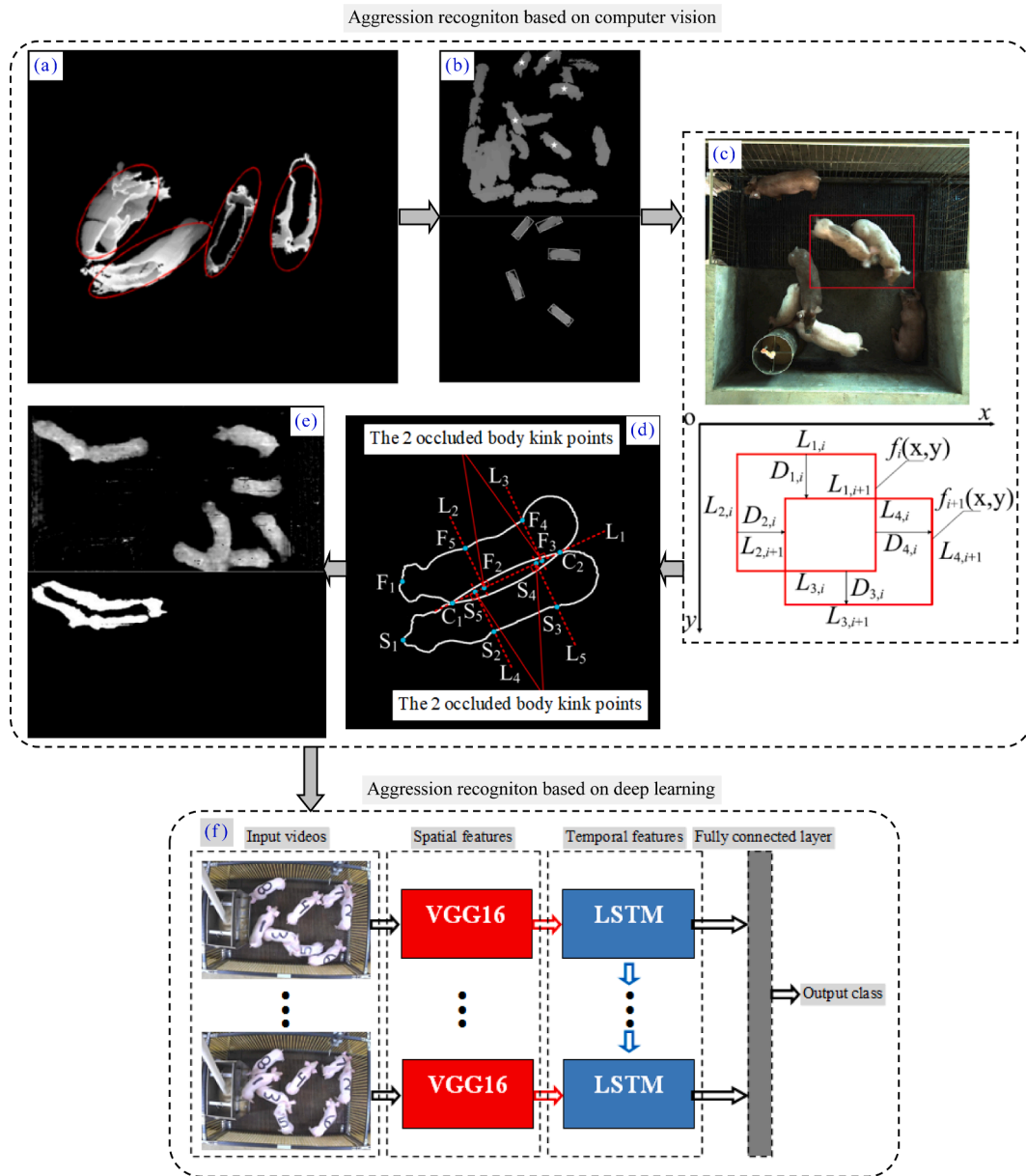


Fig. 3. Aggressive behaviour recognition based on computer vision and deep learning.

motion features on pig body from the aforementioned rectangle, Chen et al. (2018) located 5 feature points on the pig contour and calculated kinetic energy difference as features by motion analysis of these 2 aggressive pigs (Fig. 3(d)). Hierarchical clustering was used to classify these features to obtain the thresholds of high and medium aggression. Based on these thresholds, rules for aggression recognition were designed to recognise medium and high aggression in each minimum recognition unit (MRU). In order to reduce the computation amount of the algorithm by setting frame-to-frame distance, Chen et al. (2019) set frame-to-frame distance and used frame difference method to obtain moving pixels. Then, the moving pixels caused by non-aggressive behaviour were removed by setting threshold of connected area (Fig. 3(e)). The number of filtered moving pixels were summed and defined as motion shape index (MSI) in each frame. After dividing all videos into 3 s-units, the maximum, mean, variance and standard deviation of MSI in each unit were extracted as features. Finally, support vector machine (SVM) was used to classify these features to recognise aggression.

The above studies were focused on extraction of spatial features of

pigs in each frame. In order to directly use spatial–temporal features in each episode to recognise aggression, Chen et al. (2020a) developed a deep learning method based on convolutional neural network (CNN) and long short-term memory (LSTM) to recognise aggressive episodes of pigs (Fig. 3(f)). First, the CNN architecture VGG16 was used to extract spatial features. These features were then input into LSTM framework to extract spatial–temporal features. Through fully connected layer, the prediction function Softmax was finally used to determine if the current episode is aggression or non-aggression. This study further improved the accuracy of aggression recognition by changing the episode length and the frame rate. In the above studies, it can be seen that recognition of aggressive behaviour of pigs has developed from computer vision to deep learning.

4.2. Drinking behaviour recognition

Studies of drinking recognition for pigs have mainly involved with single-pig drinking recognition and multi-pig drinking recognition. In order to recognise drinking of single-pig in a farrowing crate, Lao et al.

(2016) located the sow head in the depth image and then recognised the sow's drinking during standing and sitting postures based on the distance between the head and drinker (Fig. 4(a)). Leonard et al. (2019) also studied this type of situation and developed a method to detect the presence of a sow's nose near the drinker. They first quantified the pixel number in a defined region close to the drinker and a drinking event was classified when the number of "nose" pixels in this region was greater than a threshold. In both of these cases, the drinking behaviour was detected by a simple position relationship between pig nose and drinker when the pig was in a space-limited environment around the drinker (Fig. 4(b)). In order to further recognise drinking of single-pig in an unlimited space, Yang et al. (2020) firstly differentiated the sow from the herd including the sow and its piglets and then extracted as spatial features the distance from the snout to the drinking nipple and head circularity for distinguishing head-up/head-down using a Fully Convolutional Network (FCN). They further defined the motion intensity of the head as temporal features using optical flow vectors and combined spatial and temporal features in the final recognition algorithm (Fig. 4(c)). From the above studies, it can be found that single-pig drinking recognition has developed from computer vision to deep learning.

In order to realise multi-pig drinking recognition, Kashiha et al. (2013) located the pig's nose relative to its body using the distance between the centroid and contour of the pig's body. When the distance between nose and drinker was less than 10 pixels and the duration of the visit to the drinker was more than 2 s, the pig was classified as drinking and the water consumption was estimated (Fig. 4(d)). Zhu et al. (2017) extracted the color moment, area and perimeter of the pig in the drinking region as the feature vectors and then calculated Euclidean distance between these vectors and a standard drinking template. This distance and the duration were used to recognise the pigs' presence and

to classify whether drinking had occurred or not (Fig. 4(e)). Zhang et al. (2019b) proposed a sow behaviour detection algorithm based on deep learning (SBDA-DL) by using the MobileNet classification network with a depth separable convolution operation. The algorithm was trained based on data from four sows to detect their drinking, urination and mounting behaviours (Fig. 4(f)). In the above studies, multi-pig drinking recognition mainly involves spatial features. However, drinking and drinker-playing can not be easily distinguished only by using spatial features extracted from image frames, as the distances from the snout to the drinker are all close in the process of drinking or drinker-playing. In order to further classify the drinking and drinker-playing of pigs in video episodes, Chen et al. (2020b) extracted spatial-temporal features by using the CNN architecture ResNet50 and LSTM to determine if the current episode is drinking or drinker-playing (Fig. 4(g)). In this study, shortening the drinking region of interest from body region to head region can improve the classification accuracy. From the above studies, it can be found that multi-pig drinking recognition has also developed from computer vision to deep learning.

4.3. Feeding behaviour recognition

The current literature on recognition of pig feeding behaviour has indicates that research has largely been focused on sows either at an individual or group level. For individual sows, Lao et al. (2016) developed a method to detect feeding behaviour with a depth sensor. This algorithm first obtained the coordinates of the body parts and feeder, and when the height of the head was found to be lower than that of the feeder and the distance between head and shoulder smaller than that of the standing sow, the sow was recognised as feeding (Fig. 5(a)). Leonard et al. (2019) developed a different depth sensor-based image processing

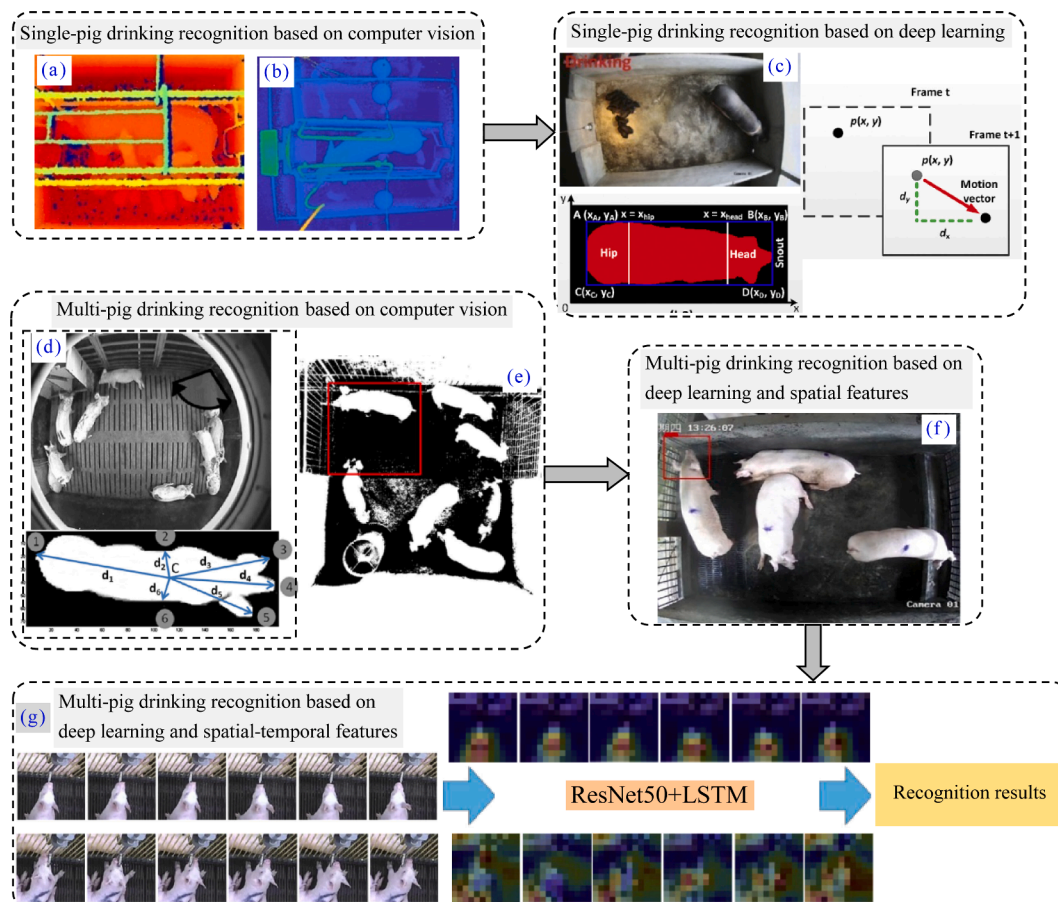


Fig. 4. Drinking behaviour recognition based on computer vision and deep learning.

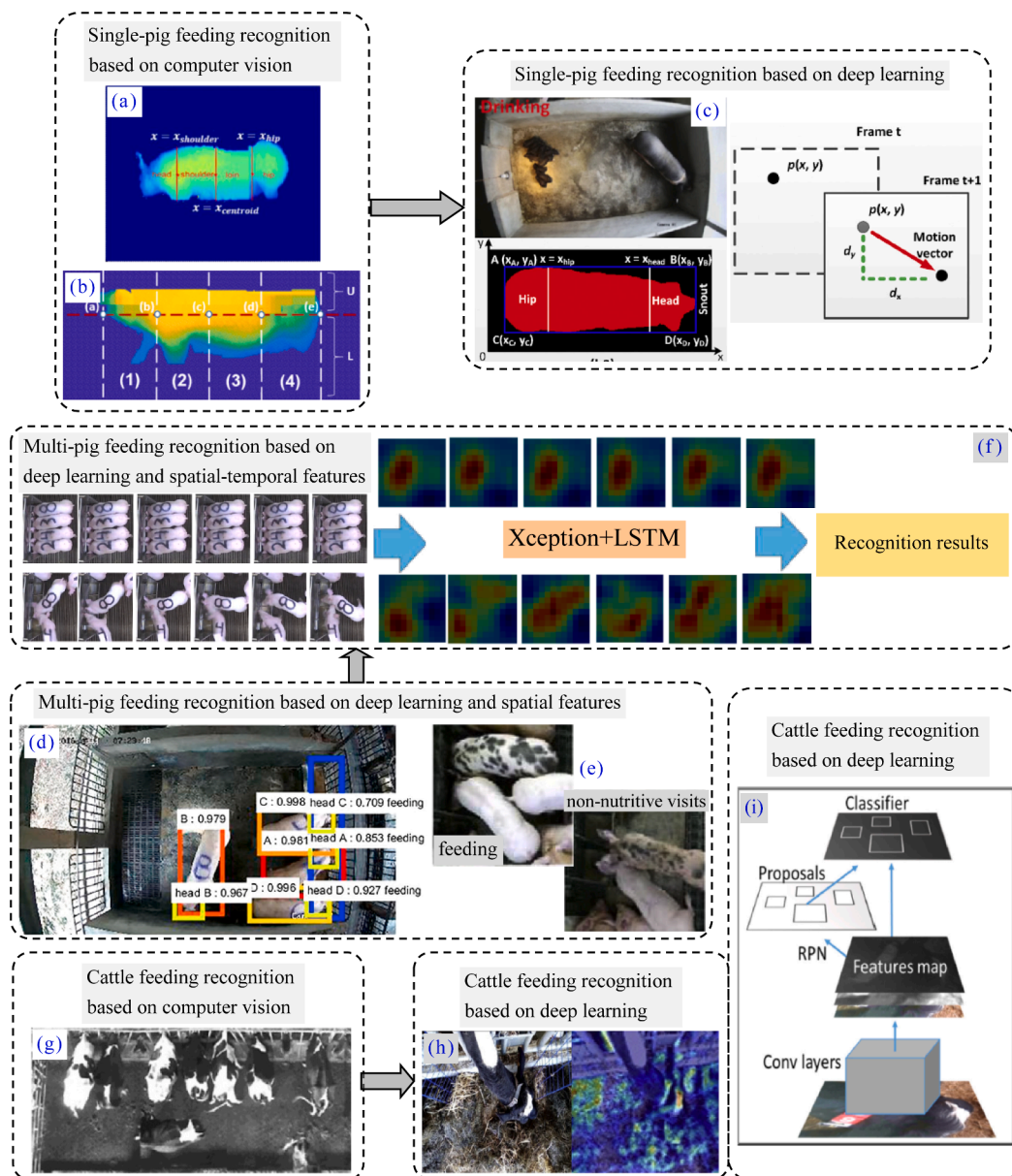


Fig. 5. Feeding behaviour recognition based on computer vision and deep learning.

method to determine a feeding sow based on the area of the sow’s head captured within a pre-determined region around the feeder. In both of the above studies the sows were constrained in farrowing stalls, where the spatial relationship between the head and feeder is small enough to allow for recognition of feeding by a single sow (Fig. 5(b)). As the behaviour of loose sows in groups will not likely be restricted by space, the recognition accuracy of the above methods may be reduced. To overcome this, Yang et al. (2020) firstly differentiated the sow from the herd including the sow and its piglets and then extracted as spatial features the circularity of the head and overlapping area of the head and feeding region using a Fully Convolutional Network (FCN). They further defined the motion intensity of the head as temporal features using optical flow vectors and combined spatial and temporal features in the final detection algorithm (Fig. 5(c)).

For feeding recognition of multiple pigs, Yang et al. (2018b) used a Faster Region-Convolutional Neural Network (Faster R-CNN) to recognise the identity of 4 sows and extracted an occupation index (i.e. ratio of the head to feeding region) as spatial features to recognise feeding behaviour of each pig (Fig. 5(d)). Alameer et al. (2020) developed a

GoogLeNet-based deep learning method for feeding detection that does not rely on pig tracking and is capable of distinguishing between feeding and non-nutritive visits (NNV) for a group of pigs (Fig. 5(e)). Chen et al. (2020c) further extracted spatial-temporal features by using the CNN architecture Xception and LSTM to recognise feeding behaviour of pigs on group level. To convert the recognition results from group level to individual level, an image processing algorithm based on maximum entropy segmentation, HSV (Hue, Saturation and Value) colour space transformation and template matching was proposed to calculate the circularity of the head, the ratio of the head to the feeding sub-region, the accumulated pixels of the head motion, and the distance from the head to the number on pig back in order to determine the identity and feeding time of each pig (Fig. 5(f)). From the above studies, it can be found that either single-pig or multi-pig feeding recognition has developed from computer vision to deep learning.

Furthermore, studies of feeding recognition for cattle have also developed from computer vision to deep learning. Porto et al. (2015) recognised cow feeding and standing behaviour by defining a methodology based on the Viola-Jones algorithm and using a multi-camera

video recording system to obtain panoramic top-view images of an area of the barn (Fig. 5(g)). On the basis of these research outcomes, the proposed system is suitable for computing cow behavioural indices and the real-time detection of behavioural changes. Achour et al. (2020) used the top of the dairy cow’s head image as a Region of Interest (ROI) and used different classifiers based on Convolutional Neural Network (CNN) model to recognise the feeding behaviour and perform individual identification of seventeen Holstein dairy cows (Fig. 5(h)). Bezen et al. (2020) designed a computer vision system for cow individual feed intake measurement based on deep Convolutional Neural Networks (CNNs) models and a low-cost RGB-D (Red, Green, Blue, Depth) camera (Fig. 5(i)). Feed intake was estimated by combining information from the RGB and depth images. Cow identification was conducted using the RGB image. Deep learning algorithms for identification and intake estimation were developed using CNN models. The training analysis shows that the model based on RGB-D data shows better results than the model based on depth channel data without RGB.

4.4. Lameness recognition

Initially, Song et al. (2008) extracted the trackway defined as “hind

hoof compared to fore hoof position” using computer vision technology after background subtraction, binary image operations, calibration and hoof separation, and then scored cow’s locomotion in order to estimate the lameness in the herd (Fig. 6(a)). Poursaberi et al. (2010) proposed a hierarchy background/foreground exaggeration to segment the cow in each frame and track it in video. The combination of logarithm and exponential, background subtraction as well as statistical filtering were used to find the accurate shape of the cow. Furthermore, the back posture of each cow during standing and walking was extracted automatically. It was done by detecting the arc of back posture and fitting a circle through selected points on the spine line. The average inverse radius of four frames displaying the hind hoofs in contact with the ground (two frames for each hoof in a row) was assigned to the cow. Based on this curvature value, a score representing the status of lameness in the individual cow was given automatically (Fig. 6(b)). The above studies apply thresholds to the whole population to detect whether or not an individual cow is lame. Viazzi et al. (2013) further developed an individualized version of the body movement pattern score, which uses back posture to classify lameness into 3 classes, and to compare both the population and the individual approach under farm conditions (Fig. 6(c)). In the above studies, two-dimensional cameras

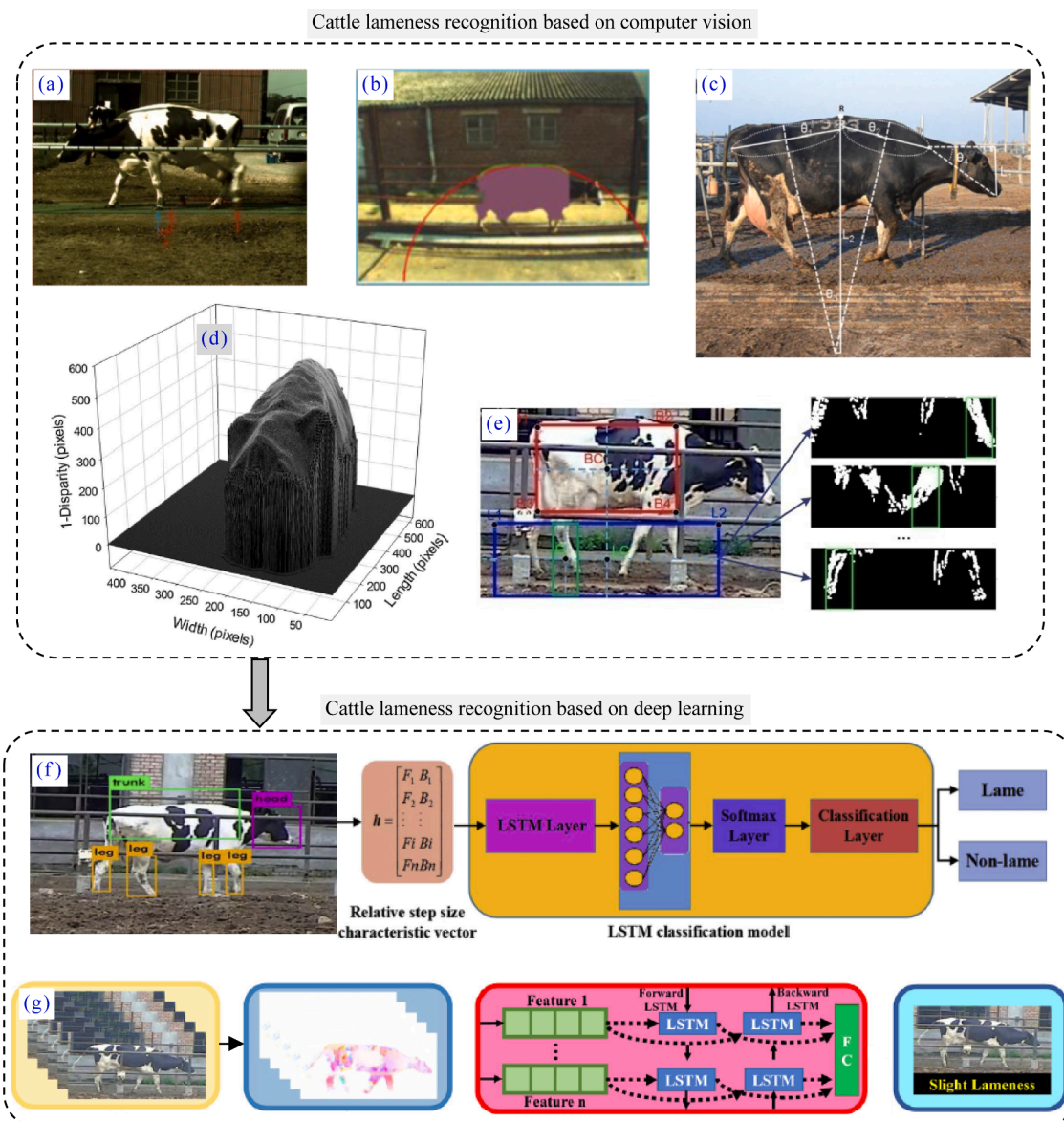


Fig. 6. Lameness recognition based on computer vision and deep learning.

positioned in side view are used to measure back posture. This method, however, is not always applicable in farm conditions since it can be difficult to be installed. Shadows and continuous changes in the background also render image segmentation difficult and often erroneous. In order to overcome these problems, [Viazzi et al. \(2014b\)](#) proposed a new computer vision method to extract the back posture from walking cows by using a three dimensional camera from top view perspective (Fig. 6 (d)). The results show that the application of a three-dimensional camera leads to an accuracy comparable to the side view approach and that the top view approach can overcome limitations in terms of automation and processing time. [Herter et al. \(2014\)](#) optimised the classification output of a computer vision-based algorithm for automated lameness scoring. A live on-the-spot assessed 5-point locomotion score using a 3D-camera was the reference for the automatic lameness score evaluation. The use of cow individual consecutive measurements improved the correct classification rate of an automatic lameness detection system. [Zhao et al. \(2018\)](#) analysed leg swing using computer vision techniques and developed an automatic and continuous system for scoring the locomotion of cows to detect and predict lameness with high accuracy and practicability (Fig. 6(e)). The motion curve was plotted by extracting the position of the moving leg by image processing, and the motion curve was analysed to generate six features referring to the gait asymmetry, speed, tracking up, stance time, stride length, and tenderness. A box-plot of the features within 3 classes showed that the dataset was nearly linear and separable under the six features and that the cows had different lameness indicators in different lameness stages.

With the development of deep learning technology, lameness recognition for cattle has developed from computer vision to deep learning. [Wu et al. \(2020\)](#) proposed a method based on YOLOv3 deep

learning algorithm and relative step size characteristic vector to classify lame and non-lame cows (Fig. 6(f)). Videos were decomposed into sequence frames, and leg targets of cows in each frame were detected by YOLOv3 algorithm. Relative step sizes of cow's front and rear legs were calculated based on leg coordinates, and the relative step size characteristic vector was constructed. Finally, a trained Long Short-Term Memory (LSTM) classification model was used to classify lame and non-lame cows based on the characteristic vector. A total of 210 videos were selected for verification using LSTM, support vector machine (SVM), K-Nearest Neighbour (KNN) and decision tree classifier (DTC) algorithms. Results showed that accuracy of lameness detection based on LSTM was 98.57%, which was 2.93%, 3.88%, and 9.25% higher than SVM, KNN, and DTC, respectively. As typical dairy cow lameness actions are only several seconds in duration and exhibit characteristic spatio-temporal structures, [Jiang et al. \(2020\)](#) attempted to capture this structure and learned action representations with convolutional neural networks (Fig. 6(g)). However, such representations are typically learned at the level of a few video frames, thereby failing to model actions over their full temporal extent. In this study, they learned video representations using neural networks with single-stream long-term optical flow convolution. The experimental results demonstrate that single-stream long-term optical flow convolution network models with increased temporal extents improve the accuracy of dairy cow lameness action recognition.

4.5. Mounting behaviour recognition

Initially, [Nasirahmadi et al. \(2016\)](#) developed a method for automatic detection of mounting events amongst pigs under commercial

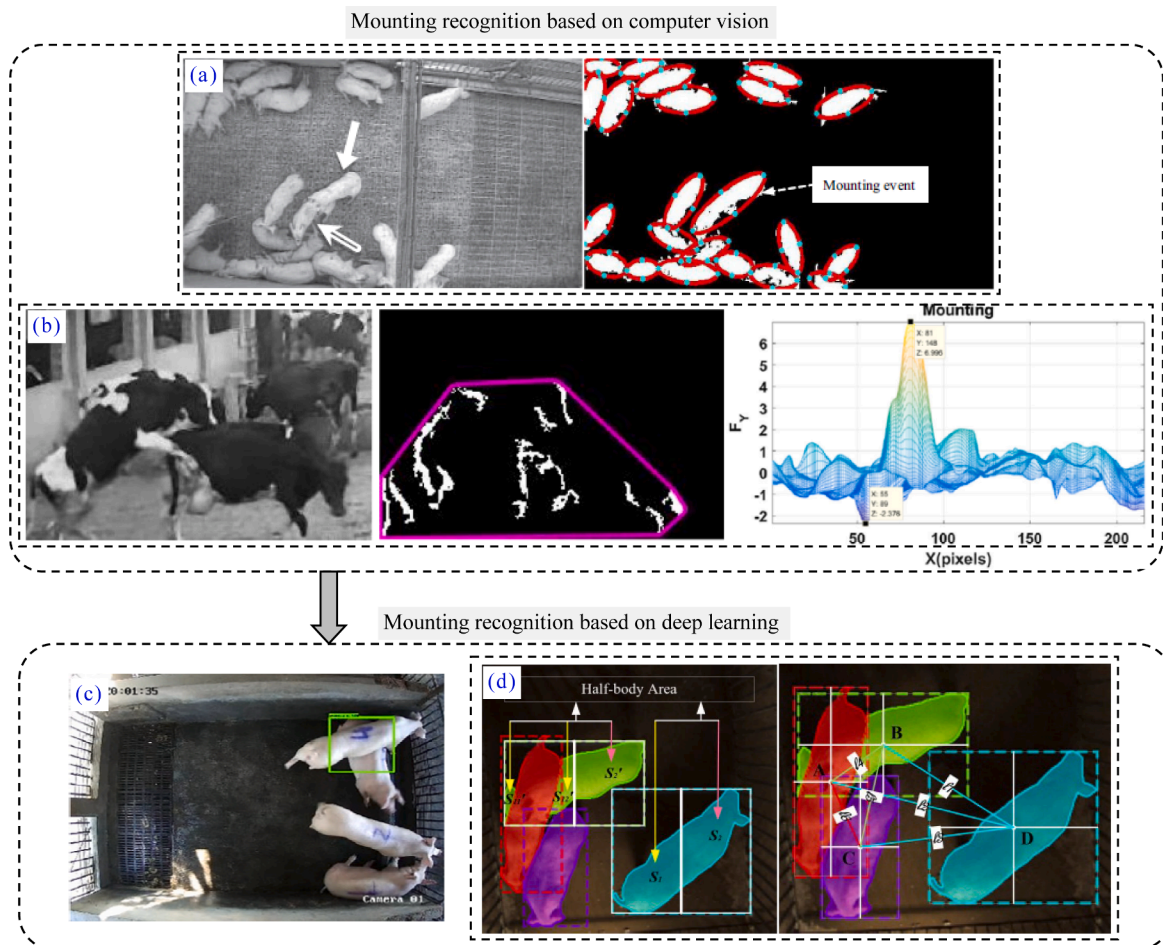


Fig. 7. Mounting behaviour recognition based on computer vision and deep learning.

farm conditions by means of image processing (Fig. 7(a)). An ellipse fitting technique was applied to localise pigs in the image. The intersection points between the major and minor axis of each fitted ellipse and the ellipse shape were used for defining the head, tail and sides of each pig. The Euclidean distances between head and tail, head and sides, the major and minor axis length of the fitted ellipse during the mounting were used for development of an algorithm to automatically recognise a mounting event. Guo et al. (2019) proposed a method for detecting mounting behavior in dairy cattle using the geometric and optical flow characteristics of identified image regions in videos taken on dairy farms (Fig. 7(b)). Firstly, masking technology was used to remove the unrelated background, convert the RGB color space to HSV color space, and adjust the summation coefficients of the HSV channels to improve the contrast between the cows and background images. Subsequently, the proposed Background Subtraction with Color and Texture Features (BSCTF) algorithm was used to detect cow regions. Then, to perform inter-frame differential processing on detection regions, the geometric and optical flow characteristics of the regions were extracted, and seven optimised features were used to construct regional feature vectors. Finally, a support vector machine (SVM) classifier was trained to classify the detected regions into mounting regions and non-mounting regions, which allowed the recognition of mounting behaviour. In the above studies, computer vision technologies were used to recognise mounting behaviour of pigs and cattle.

Furthermore, the studies of mounting recognition for pigs have developed from computer vision to deep learning. Zhang et al. (2019b) proposed a Real-Time Sow Behavior Detection Algorithm based on Deep Learning (SBDA-DL) to recognise pigs' mounting, drinking and urination (Fig. 7(c)). The algorithm uses an optimised deep learning network structure to directly detect the sow behavior. This improves the accuracy of behavior detection at a processing speed required for real-time detection and meets the requirements of daily monitoring from auxiliary staff in most pig breeding farms. Li et al. (2019b) developed an efficient learning algorithm to recognise mounting behaviour of pigs based on the data characteristics of visible light images (Fig. 7(d)). Pig segmentation network based on Mask Region-Convolutional Neural Networks (Mask R-CNN) was applied to extract individual pigs in the frames. The region of interest (RoI) parameters and mask coordinates of each pig, from which eigenvectors were extracted, could be obtained. Subsequently, the eigenvectors were classified with a kernel extreme learning machine (KELM) to determine whether mounting behaviour has occurred. This method can be an efficient way of solving the problem of segmentation difficulty caused by partial occlusion and adhesion of pig bodies, even if the pig body colour was similar to the background, in recognition of mounting behaviour.

4.6. Behaviour posture recognition

For recognition of behaviour posture in pigs, Nasirahmadi et al. (2015) investigated the feasibility of using image processing and the Delaunay triangulation method to detect change in group lying behaviour of pigs under commercial farm conditions and relate this to changing environmental temperature (Fig. 8(a)). Nasirahmadi et al. (2019a) investigated whether a two-dimensional imaging system could be used for lateral and sternal lying posture detection in pigs under commercial farm conditions (Fig. 8(b)). Pigs were extracted from their background using a background subtracting method. Based on the binary image properties, the boundaries and convex hull of each animal were found. In order to determine their lying posture, the area and perimeter of each boundary and convex hull were calculated in lateral and sternal lying postures as inputs for training of a linear SVM classifier. The trained SVM was then used to detect the target postures in binary images. In above studies, computer vision technologies were used for recognition of pig postures.

Recognition of pig postures has developed from computer vision to deep learning. Zheng et al. (2018) introduced a detector, Faster R-CNN,

on deep learning framework to identify five postures (standing, sitting, sterna recumbency, ventral recumbency and lateral recumbency) and obtain sows accurate location in loose pens (Fig. 8(c)). The detection system consists of a Kinect v2 sensor that acquires depth images and a program that identifies sow postures and locates its bounding-boxes. Nasirahmadi et al. (2019b) determined whether a two-dimensional imaging system, along with deep learning approaches, could be utilised to detect the standing and lying (belly and side) postures of pigs under commercial farm conditions (Fig. 8(d)). Three deep learning-based detector methods, including faster regions with convolutional neural network features (Faster R-CNN), single shot multibox detector (SSD) and region-based fully convolutional network (R-FCN), combined with Inception V2, Residual Network (ResNet) and Inception ResNet V2 feature extractions of RGB images were proposed. Riekert et al. (2020) designed a deep learning system for position and posture detection that only requires standard 2D camera imaging with no adaptations to the application setting (Fig. 8(e)). This deep learning system applies the state-of-the-art Faster R-CNN object detection pipeline and the state-of-the-art Neural Architecture Search (NAS) base network for feature extraction. Zhu et al. (2020) proposed an end-to-end refined two-stream RGB-D Faster region convolutional neural network (R-CNN) algorithm, which fuses RGB-D image features in the feature extraction stage for recognising five postures of lactating sows (standing, sitting, sternal recumbency, ventral recumbency, and lateral recumbency) in scenes at a pig farm (Fig. 8(f)). Based on the Faster R-CNN algorithm, two CNNs were first used to extract the RGB image features and depth image features. Then, a proposed single RGB-D region proposal network was used to generate the regions of interest (ROIs) for the two types of image feature maps in RGB-D. Next, the features of the RGB-D ROIs were extracted and merged using a feature fusion layer. Finally, the fused features of the RGB-D ROIs were input into a Fast R-CNN to obtain the recognition results. In order to detect frame-level sow postures, temporally localise posture change actions, and generate spatio-temporally action tubes parsed from a long-time untrimmed segment of depth video, Zheng et al. (2020) recorded depth videos for five batches of lactating sows, using a Kinect from a top-view in a commercial farm. Three batches were used for training and validation, and the other two for testing. Four postures (standing, sitting, ventral lying, and lateral lying) were automatically detected (Fig. 8(g)). The methodology can be applied in large-scale deployments for learning livestock action preferences and behavioural traits, thereby enhancing welfare and productivity on a farm.

For recognition of behaviour posture in cattle, Porto et al. (2013) proposed a computer vision-based system for the automatic detection of dairy cow lying behaviour in free-stall barns (Fig. 8(h)). The system was composed of a multi-camera video-recording system and a software component which executed a cow lying behaviour detector model using the Viola-Jones algorithm. The experimental results suggest that the system proposed in this study could be used for the calculation of the cow lying index which is widely used to investigate cow lying behaviour in free-stall barns. Li et al. (2019c) developed three deep cascaded convolutional neural networks models, including the convolutional pose machine model, the stacked hourglass model and the convolutional heatmap regression model to perform robust cattle pose estimation, with RGB images captured under real cattle farm conditions (Fig. 8(i)). From the above studies, it can be found that recognition of behaviour posture in cattle has also developed from computer vision to deep learning.

4.7. Tail-biting behaviour recognition

Presently, studies of tail-biting recognition for pigs have been few. Liu et al. (2020b) developed a deep learning method to identify and locate tail-biting interactions in pigs. This method used a tracking-by-detection algorithm to simplify the group-level behaviour to pairwise interactions. Then, a convolution neural network (CNN) and a recurrent neural network (RNN) were combined to extract the spatial-temporal

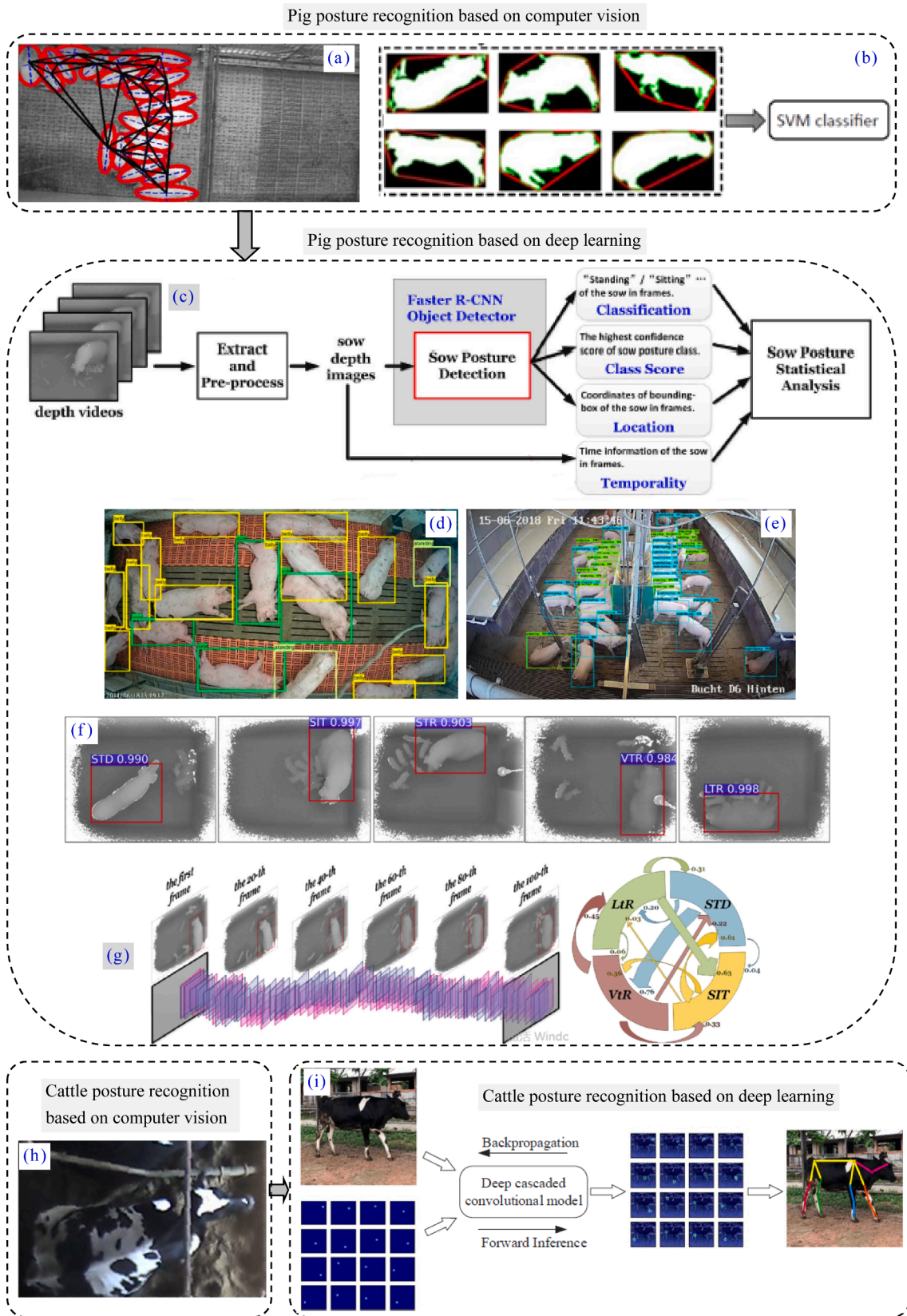


Fig. 8. Behaviour posture recognition based on computer vision and deep learning.

features to recognise tail-biting of pigs (Fig. 9). The results demonstrate that the tracking-by-detection approach is capable of obtaining the trajectories of biters and victims.

4.8. Nursing behaviour recognition

Initially, Yang et al. (2018c) developed a computer vision-based method for automatic recognition of nursing interactions under commercial farm conditions by using spatial and temporal information of nursing behaviour (Fig. 10(a)). For spatial information extraction, the spatial distribution between the mother sow and her piglets during nursing was used to detect the possible nursing episodes. Sows were segmented accurately by a fully convolutional network, and udder zones were calculated dynamically by the geometrical properties of the nursing sow and the piglet length. Spatial information from piglets was extracted in the self-adaptive udder zones. For temporal information extraction, to distinguish behaviours similar to nursing, temporal motion information about the intensity of motion and the occupation index was extracted from optical flow of the udder zones. The results show that the recognition method designed represents a significant step forward in automatically identifying nursing behaviour in commercial pig farms. Subsequently, Yang et al. (2019) developed a method to automatically recognise nursing behaviors of sow in videos by exploiting the spatio-temporal relations. The method firstly detected spatio-temporal key cuboids which may have nursing interactions in them, and then Oriented Nursing Flows (ONuF) of the spatio-temporal key cuboids was proposed to further recognise nursing behaviours (Fig. 10(b)). In the first step, the temporal key episodes were first detected in video using optical flow-based features containing the distribution of nursing pixel and distance among moving pixels to estimate the distribution of motion all over the frame. Then, spatial key regions in the temporal key episodes were located by identifying the spatial position and geometric properties of the sow and her piglets using a fully convolutional network-based semantic segmentation approach. In the second step, to further recognise nursing behavior, a new feature descriptor ONuF, estimating the motion orientation change and motion magnitude of spatio-temporal key cuboids, was proposed and used to learn a SVM classifier. The

results indicate that the method exploiting the spatio-temporal relations using fully convolutional network and oriented optical flow can be used for automatically recognising nursing behaviours from daily behavioural videos of lactating sows.

4.9. Playing behaviour recognition

In recent literature, there is few study of playing recognition for pigs. Chen et al. (2020d) developed a computer vision based approach that utilised a recurrent neural network-based deep learning algorithm to recognise pig playing behaviours and preliminarily determine the preference to objects. Firstly, the HSV (Hue, Saturation, Value) colour space-based tracking algorithm was developed to locate object region of interest. Secondly, the convolutional neural network (CNN) architecture InceptionV3 was used to extract spatial features from each frame. These features were input into the long short-term memory (LSTM) framework to extract spatial-temporal features from each episode. Through the fully connected layer, the prediction function Softmax was finally used to classify these episodes as playing or non-playing (Fig. 11). These results indicate that the proposed method can be used to recognise playing behaviours of pigs, and halving the radius of the region of interest can improve the recognition accuracy of playing behaviours. Moreover, the preference of pigs to objects based on playing duration can be determined. The obtained duration of playing behaviours can help farmers to evaluate the enrichment used and thereby to increase the health and welfare of the pigs in their care.

4.10. Multi-behaviour recognition

Initially, Lao et al. (2016) developed a computer vision-based system that automatically recognises sow behaviors (i.e., lying, sitting, standing, kneeling, feeding, drinking, and shifting) in farrowing crate. The system consisted of a low-cost 3D camera that simultaneously acquires digital and depth images and a software program that detects and identifies the sow's behaviors (Fig. 12(a)). This study describes the computational algorithm for the analysis of depth images and presents its performance in recognising the sow's behaviours as compared to

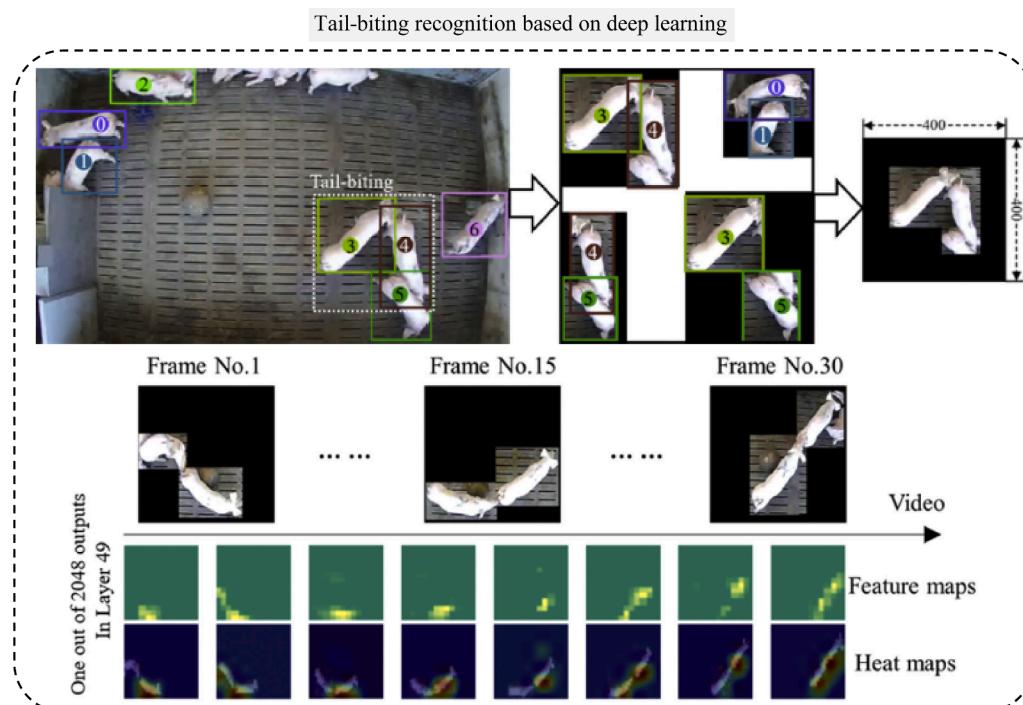


Fig. 9. Tail-biting behaviour recognition based on computer vision and deep learning.

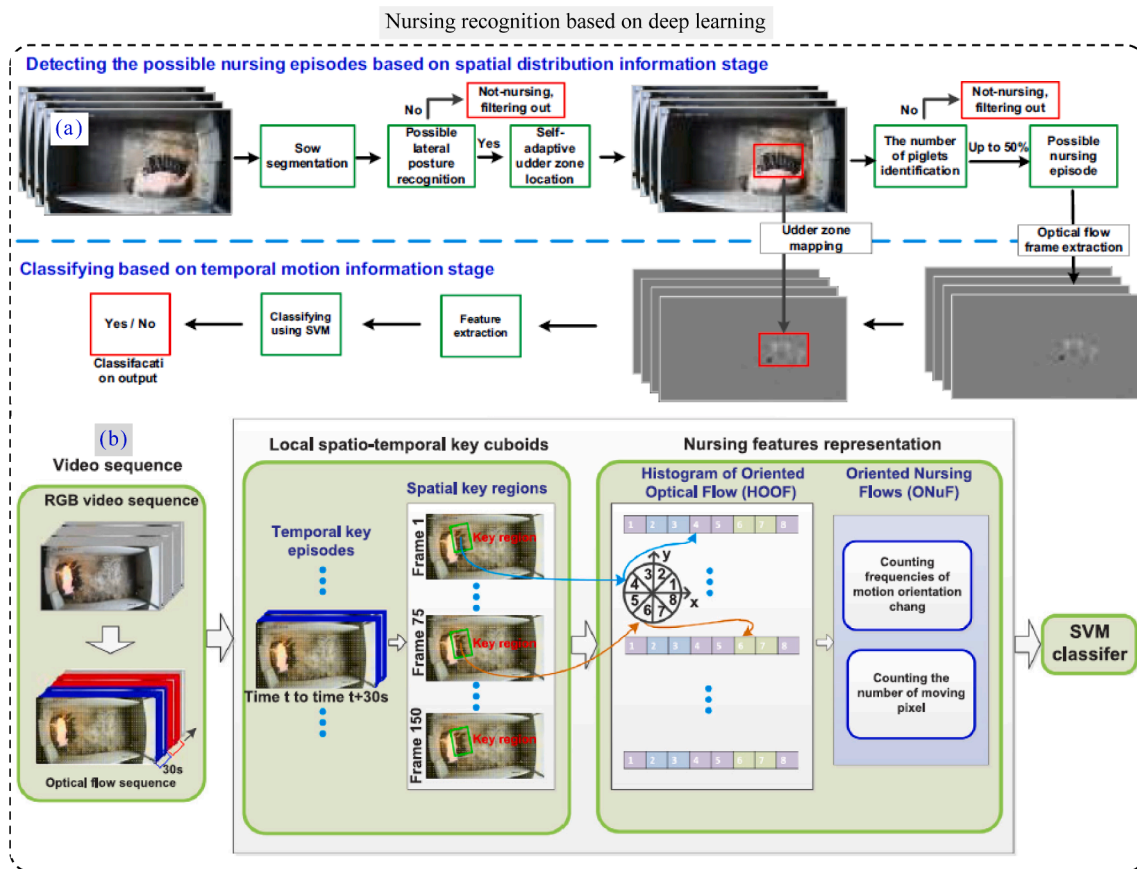


Fig. 10. Nursing behaviour recognition based on computer vision and deep learning.

manual recognition. Leonard et al. (2019) monitored the behaviours of sows and piglets in a commercial setting utilising an autonomous machine vision system (Fig. 12(b)). The objectives of this research were to implement a digital and time-of-flight depth imaging system, develop a process with minimal user input to analyse the collected images, and calculate the hourly and daily posture and behaviour budgets of sows housed in individual farrowing stalls. This autonomous system enables acquisition of a large amount of replicated data to evaluate the effects of changing the farrowing environment on sow behavior and potentially well-being. The above studies indicate that computer vision is enough for recognition of multiple behaviours of pigs in a limited space.

For unlimited space, deep learning has been used for recognition of multiple behaviours of pigs. Yang et al. (2020) developed a general and automatic recognition framework for recognising the daily behaviours of lactating sows to save manual labour and promote smart management. The proposed framework used both image analysis techniques in still image and motion analysis techniques in spatiotemporal videos to recognise sow drinking, feeding, nursing, moving, medium active and inactive behaviours in a loose pen (Fig. 12(c)). The image analysis techniques, which are based on fully convolutional networks (FCNs) for high-accuracy segmentation, were used to extract spatial features that evaluated the spatial relationships between objects and the appearance of sows. The motion analysis techniques in spatiotemporal videos, which are based on optical flow analysis and changes in the animal centroid, were used to extract temporal features that evaluated the temporal motions of the animals. In the recognition process, these spatial and temporal features were input into a hierarchical classifier for behaviour recognition. Recognition results were obtained by a temporal-correlation-based correction module for promoting the recognition rate. The proposed method provides an effective approach for the automatic recognition of sow behaviours from video sequences, which

facilitates the pig farmer in improving livestock-farming management. Zhang et al. (2019b) proposed a Real-Time Sow Behavior Detection Algorithm based on Deep Learning (SBDA-DL) by using the MobileNet classification network with a depth separable convolution operation. The algorithm was used for real-time detection of three typical sow behaviours, i.e., drinking, urination, and mounting (Fig. 12(d)). Zhang et al. (2020) took image frames and optical flow from videos as two-stream input objects to fully extract the temporal and spatial behavioural characteristics. Two-stream convolutional network models based on deep learning were proposed, including inflated 3D convnet (I3D) and temporal segment networks (TSN) whose feature extraction network is Residual Network (ResNet) or the Inception architecture (e.g., Inception with Batch Normalization (BN-Inception), InceptionV3, InceptionV4, or InceptionResNetV2) to achieve recognition of pig behaviours, i.e., feeding, lying, walking, scratching and mounting (Fig. 12(e)). The experimental results indicated that The TSN model (ResNet101) is superior to the other models in solving the task of pig behavior recognition. Li et al. (2020) proposed a pig behaviour recognition network with a spatiotemporal convolutional network based on the SlowFast network architecture for behaviour classification of five categories (Fig. 12(f)). Firstly, a pig behavior recognition video dataset (PBVD-5) was built by cutting short clips from 3-month non-stop shooting videos, which was composed of five categories of pig's behaviour: feeding, lying, motoring, scratching and mounting. Subsequently, a SlowFast network based spatiotemporal convolutional network for the pig's multi-behavior recognition (PMB-SCN) was proposed. The results of the networks with variant architectures of the PMB-SCN were implemented and the optimal architecture was compared with the state-of-the-art single stream 3D convolutional network in the dataset. The results showed that this network provided remarkable ability of generalization and possibility for the subsequent pig detection

Playing recognition based on deep learning

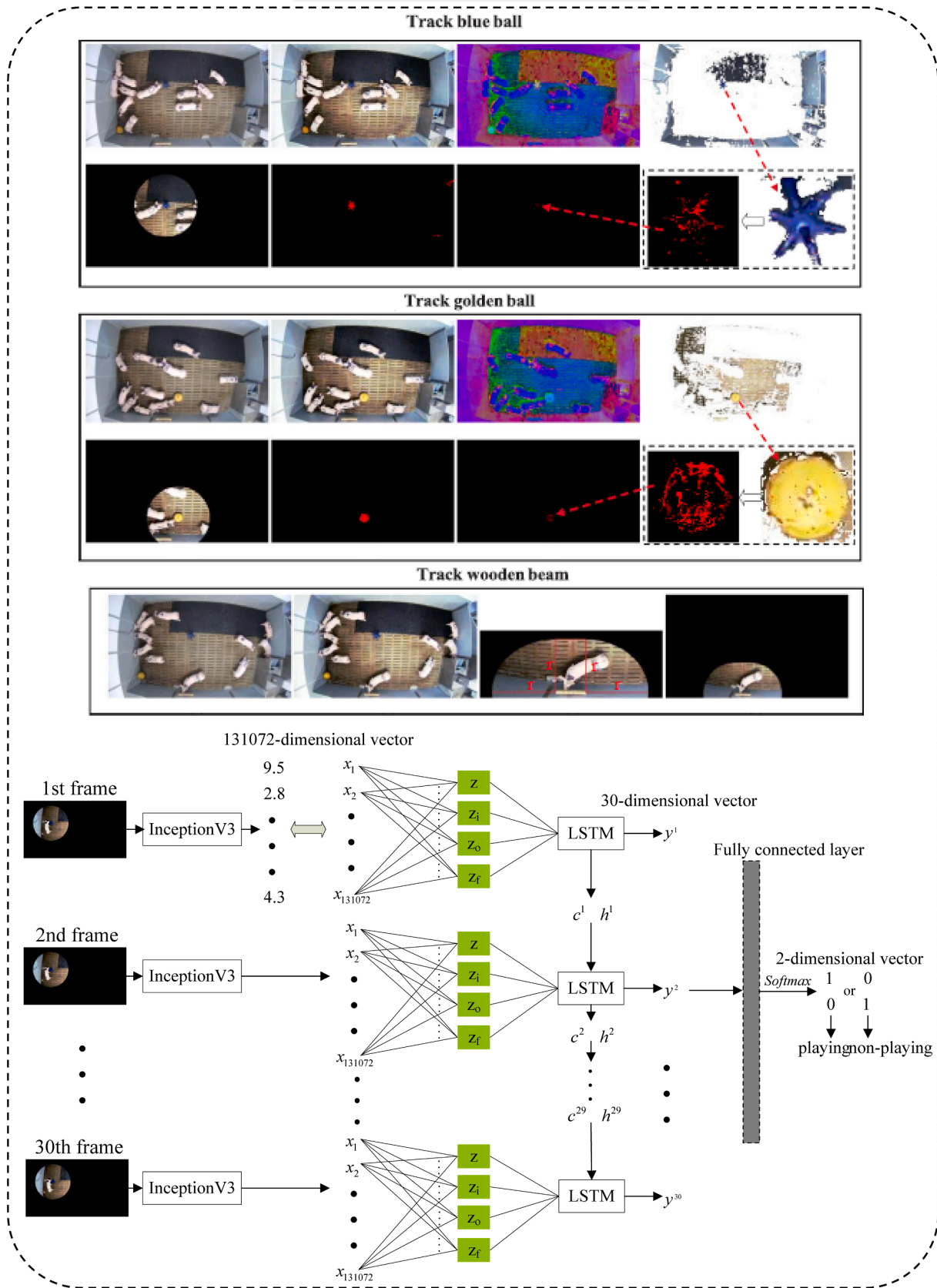


Fig. 11. Playing behaviour recognition based on computer vision and deep learning.

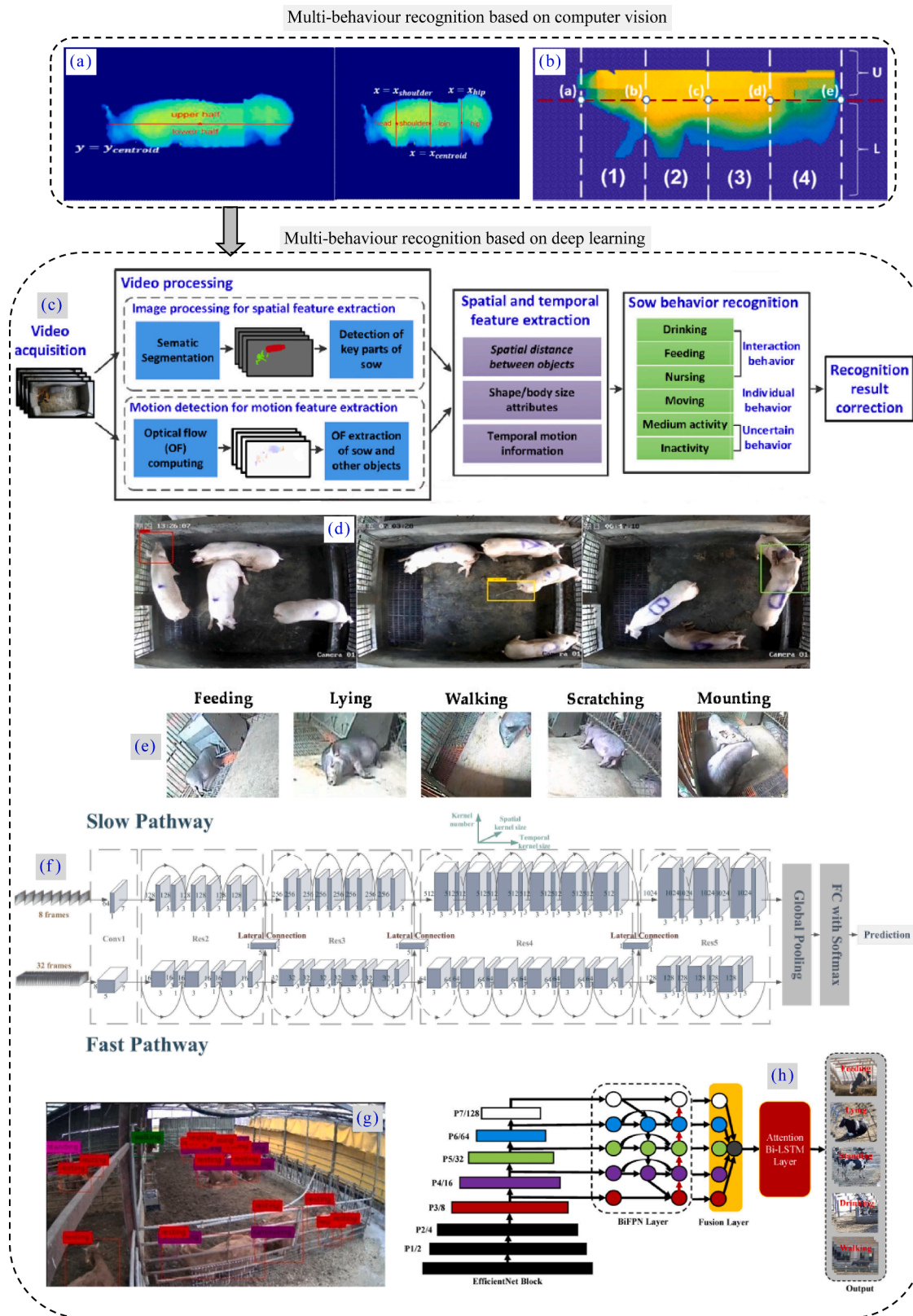


Fig. 12. Multi-behaviour recognition based on computer vision and deep learning.

and behaviour recognition simultaneously.

On the other hand, deep learning has also been used for recognition of multiple behaviours of cattle. Fuentes et al. (2020) introduced an approach for hierarchical cattle behavior recognition with spatial-temporal information based on deep learning (Fig. 12(g)). This

framework involved appearance features at frame-level and spatio-temporal information that incorporates more context-temporal features. The system can detect (class) and localize (bounding box) regions containing multiple cattle behaviors in the video frames. The results show that the system can effectively recognise 15 different types of

hierarchical activities divided into individual and group activities, and also part actions. Yin et al. (2020) achieved the recognition of cows' lying, standing, walking, drinking and feeding behaviours (Fig. 12(h)). First, based on the advantage of the efficient feature extraction of EfficientNet, the spatial feature extraction of cow's video frames was realized. Then, to fully extract the characteristics of different behaviour information of dairy cows, the BiFPN (bidirectional feature pyramid network) was used to realize the efficient fusion of characteristics in the 3–5 layers of EfficientNet. Finally, the behaviour information was sent to the BiLSTM (bidirectional long short-term memory) module, which integrates the attention mechanism to realise the aggregation of video frames in a time series, thereby realising fast and accurate recognition of dairy cow's motion behaviours.

5. Research trend

5.1. Development of robust livestock identification algorithms

Livestock identification is a basic step of transforming group behaviour recognition into individual behaviour recognition. From the existing studies of identification of pigs and cattle, it can be seen that the region on animal body used for identification has moved from body part, e.g. pig face (Hansen et al., 2018; Marsot et al., 2020) and cattle muzzle (Gaber et al., 2016; Kumar et al., 2018), to the body entirety. Moreover, the technologies used have developed from computer vision to deep learning. The local region on moving animal is difficult to locate and it requires that animal enters the designated area cooperatively to obtain local images, which may involve human intervention. Therefore, future identification of livestock will develop in the overall direction of the body entirety. However, there is still a lot of space for improvement in the current identification algorithms based on body entirety. For instance, the performance of the computer vision method based on Gabor filters Local Binary Pattern (Huang et al., 2018) will be influenced due to the touching or overlapping between pigs. Furthermore, the performance of the deep learning method based on Faster R-CNN (Yang et al., 2018b) may be reduced when the number of pigs increases largely, especially in the nursery pens where the number of pigs is sometimes more than 30 in some European pig farms (Nasirahmadi et al., 2019a, 2019b). It is because that the deep learning-based object detection algorithms (e.g. the CNN architectures Faster R-CNN (Girshick, 2015), SSD (Liu et al., 2016) and YOLO (Redmon et al., 2016)) are mainly used for classification of different types of objects, while the pigs are often similar in the same pen. Even if in the case of CNN and LSTM architecture for identification of the cattle (Qiao et al., 2019b), the performance of the CNN architecture will determine the final identification results, as CNN mainly extracts the spatial features in the image while LSTM only fuses these spatial features in the temporal dimension. Therefore, the improvement of the CNN architecture has become one of the main research directions of robust livestock identification.

5.2. Recognition of livestock behaviours for different growth stages

The present studies of behaviour recognition for livestock have mainly focused on the animal's specific growth stages, e.g. lactating period (Yang et al., 2019), nursery period (Chen et al., 2020a), and fattening period (Li et al., 2020). The performance of algorithm of livestock behaviour recognition in one stage may be influenced when this algorithm is used for recognition of behaviours in another stage, as the size of the pig body and the stocking density are different. For instance, usually the number of pigs is small and the pen is loose in the fattening sow pens (Zhang et al., 2019b), while the number of pigs and the stocking density are larger in the nursery piglet pens, which will generate more occlusion, touching and overlapping between pigs when the pig is drinking (Chen et al., 2020b). Therefore, the performance of the drinking recognition algorithm developed in fattening pens may be reduced when it is used to recognise drinking of pigs in nursery pens. In

order to address this problem, deep learning technologies, e.g., CNN and LSTM architectures (Donahue et al., 2015; Srivastava et al., 2015), can be used to extract the spatial-temporal motion pattern of the same behaviour in different growth stages of livestock for recognition of this behaviour, as deep learning has ability to train amounts of data into a robust model. On the other hand, the trained model could be retrained if there is a great difference of the motion pattern of the same behaviour among different growth stages.

5.3. Further quantification of the results of behaviour recognition

In recent years, livestock behaviour studies based on computer vision and deep learning has mainly focused on the recognition of a specific behaviour, e.g., aggression (Chen et al., 2020a), drinking (Chen et al., 2020b), feeding (Chen et al., 2020c), lameness (Jiang et al., 2020), mounting (Li et al., 2019b), posture (Zheng et al., 2020), tail-biting (Liu et al., 2020b), nursing (Yang et al., 2020) and playing (Chen et al., 2020d), or classification of multiple behaviours, e.g. feeding, lying, walking, scratching and mounting (Zhang et al., 2020; Li et al., 2020). However, the recognition or classification results of these livestock behaviours need to be further quantified. For instance, the present studies have mainly focused on the recognition of pig feeding and the quantification of feeding time of each pig (Yang et al., 2020; Chen et al., 2020c). While feed intake, meal length, meal interval, number of eaten meals per day, and total time spent eating (Marcon et al., 2015; Adrian et al., 2018; Cross et al., 2018) are also necessary to be quantified, as changes in the quantity of feed eaten by pigs can indicate environmental, health or social stressors (Matthews et al., 2016) and can also be used to predict outbreaks of diseases such as pneumonia (Brown-Brandl et al., 2016). Therefore, further quantification of the results of behaviour recognition has become one of the important research directions of livestock behaviour recognition in the future.

5.4. Building evaluation system of growth status, health and welfare

Presently, the existing studies of livestock behaviours has mainly involved with recognition of feeding, drinking, aggression, lameness, mounting, tail-biting, posture, nursing and playing. Among them, feeding and drinking can directly reflect the growth status of livestock animals (Botreau et al., 2007). Aggression, lameness, mounting, tail-biting can cause animals' skin trauma, infection and even fatal injuries, which are closely related to the animal health (Turner et al., 2006; Bruijnjs et al., 2012; Teixeira and Boyle, 2014; Larsen et al., 2019). Pose estimation is also a key step evaluating animal health (Yazdanbakhsh et al., 2017). Furthermore, nursing and playing mainly reflect animal welfare (Muns et al., 2013; Lahrman et al., 2018). However, the existing studies of recognition of the above livestock behaviours have mainly realised the recognition of the specific behaviours. For the aspects of the evaluation of growth status, health and welfare, there are few parameters that can build a relationship between the results of behaviour recognition and the growth status, health and welfare of animals. For instance, the parameters of feeding/ drinking amount or animal weight could be quantified and related to the results of recognition of feeding and drinking in order to evaluate animal growth status. The parameter of the degree of skin trauma could be quantified and related to the results of recognition of aggression, lameness, mounting and tail-biting in order to evaluate animal health. Furthermore, the parameter of the preference to different enrichment objects could be quantified and related to the results of recognition of playing in order to evaluate animal welfare. Therefore, building evaluation system of growth status, health and welfare has been the main research direction used for establishing the relationship between the results of livestock behaviour recognition and relevant evaluation parameters, which will greatly help farmers to evaluate the growth status, health and welfare of animals.

6. Conclusion

In recent years, many studies of recognition of livestock behaviours based on computer vision and deep learning have appeared. As pigs and cattle are typical commercial livestock and their aggression, drinking, feeding, lameness, mounting, posture, tail-biting, nursing, playing and other behaviours are closely related to the growth status, health and welfare, this paper mainly evaluates the methods for recognition of these behaviours of pigs and cattle. Since image segmentation and identification are the basis of livestock behaviour recognition, this paper analyses the development process of image segmentation, identification and behaviour recognition from computer vision to deep learning, which provides researchers and producers with technical references. In addition, this paper further elaborates the research trend of livestock behaviour recognition from four aspects, i.e., development of robust livestock identification algorithms, recognition of livestock behaviours for different growth stages, further quantification of the results of behaviour recognition, and building evaluation system of growth status, health and welfare.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work was a part of the project funded by the “National Natural Science Foundation of China”, China (grant number: 31872399). Tomas Norton would like to acknowledge the support from Pig Improvement Company (PIC) for his contribution to this work.

References

- Allen, A., Golden, B., Taylor, M., Patterson, D., Henriksen, D., Skuce, R., 2008. Evaluation of retinal imaging technology for the biometric identification of bovine animals in Northern Ireland. *Livestock Sci.* 116 (1-3), 42–52.
- Adrion, F., Kapun, A., Eckert, F., Holland, E.M., Staiger, M., Götz, S., Gallmann, E., 2018. Monitoring trough visits of growing-finishing pigs with UHF-RFID. *Comput. Electron. Agric.* 144, 144–153.
- Ahrendt, P., Gregersen, T., Karstoft, H., 2011. Development of a real-time computer vision system for tracking loose-housed pigs. *Comput. Electron. Agric.* 76 (2), 169–174.
- Alameer, A., Kyriazakis, I., Dalton, H.A., Miller, A.L., Bacardit, J., 2020. Automatic recognition of feeding and foraging behaviour in pigs using deep learning. *Biosyst. Eng.* 197, 91–104.
- Achour, B., Belkadi, M., Filali, I., Laghrouche, M., Lahdir, M., 2020. Image analysis for individual identification and feeding behaviour monitoring of dairy cows based on Convolutional Neural Networks (CNN). *Biosyst. Eng.* 198, 31–49.
- Botreau, R., Veissier, I., Butterworth, A., Bracke, M.B.M., Keeling, L.J., 2007. Definition of criteria for overall assessment of animal welfare. *Anim. Welf.* 16 (2), 225–228.
- Brown-Brandl, T.M., Jones, D.D., Eigenberg, R.A., 2016. Modeling feeding behavior of swine to detect illness. 2016 CIGR-AgEng Conference. Aarhus, Denmark: CIGR.
- Brujinis, M.R.N., Beerda, B., Hogeveen, H., Stassen, E.N., 2012. Assessing the welfare impact of foot disorders in dairy cattle by a modeling approach. *Animal* 6 (6), 962–970.
- Brünger, J., Traulsen, I., Koch, R., 2018. Model-based detection of pigs in images under sub-optimal conditions. *Comput. Electron. Agric.* 152, 59–63.
- Bezen, R., Edan, Y., Halachmi, I., 2020. Computer vision system for measuring individual cow feed intake using RGB-D camera and deep learning algorithms. *Comput. Electron. Agric.* 172, 105345. <https://doi.org/10.1016/j.compag.2020.105345>.
- Chen, C., Zhu, W.X., Ma, C.H., Guo, Y.Z., Huang, W.J., Ruan, C.Z., 2017. Image motion feature extraction for recognition of aggressive behaviours among group-housed pigs. *Comput. Electron. Agric.* 142, 380–387.
- Chen, C., Zhu, W., Guo, Y., Ma, C., Huang, W., Ruan, C., 2018. A kinetic energy model based on machine vision for recognition of aggressive behaviours among group-housed pigs. *Livestock Sci.* 218, 70–78.
- Chen, C., Zhu, W., Liu, D., Steibel, J., Siegford, J., Wurtz, K., Han, J., Norton, T., 2019. Detection of aggressive behaviours in pigs using a RealSense depth sensor. *Comput. Electron. Agric.* 166, 105003. <https://doi.org/10.1016/j.compag.2019.105003>.
- Chen, C., Zhu, W., Steibel, J., Siegford, J., Wurtz, K., Han, J., Norton, T., 2020a. Recognition of aggressive episodes of pigs based on convolutional neural network and long short-term memory. *Comput. Electron. Agric.* 169, 105166. <https://doi.org/10.1016/j.compag.2019.105166>.
- Chen, C., Zhu, W., Steibel, J., Siegford, J., Han, J., Norton, T., 2020b. Classification of drinking and drinker-playing in pigs by a video-based deep learning method. *Biosyst. Eng.* 196, 1–14.
- Chen, C., Zhu, W., Steibel, J., Siegford, J., Han, J., Norton, T., 2020c. Recognition of feeding behaviour of pigs and determination of feeding time of each pig by a video-based deep learning method. *Comput. Electron. Agric.* 176, 105642. <https://doi.org/10.1016/j.compag.2020.105642>.
- Chen, C., Zhu, W., Oczak, M., Maschat, K., Baumgartner, J., Larsen, M.L.V., Norton, T., 2020d. A computer vision approach for recognition of the engagement of pigs with different enrichment objects. *Comput. Electron. Agric.* 175, 105580. <https://doi.org/10.1016/j.compag.2020.105580>.
- Donahue, J., Hendricks, L.A., Guadarrama, S., Rohrbach, M., Venugopalan, S., Saenko, K., Darrell, T., 2015. Long-term recurrent convolutional networks for visual recognition and description. In: *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 2625–2634.
- Cross, A.J., Rohrer, G.A., Brown-Brandl, T.M., Cassady, J.P., Keel, B.N., 2018. Feed-forward and generalised regression neural networks in modelling feeding behaviour of pigs in the grow-finish phase. *Biosyst. Eng.* 173, 124–133.
- Fuentes, A., Yoon, S., Park, J., Park, D.S., 2020. Deep learning-based hierarchical cattle behavior recognition with spatio-temporal information. *Comput. Electron. Agric.* 177, 105627. <https://doi.org/10.1016/j.compag.2020.105627>.
- Guo, Y., Zhu, W., Jiao, P., Chen, J., 2014. Foreground detection of group-housed pigs based on the combination of Mixture of Gaussians using prediction mechanism and threshold segmentation. *Biosyst. Eng.* 125, 98–104.
- Guo, Y.-Z., Zhu, W.-X., Jiao, P.-P., Ma, C.-H., Yang, J.-J., 2015. Multi-object extraction from topview group-housed pig images based on adaptive partitioning and multilevel thresholding segmentation. *Biosyst. Eng.* 135, 54–60.
- Girshick, R., 2015. Fast R-CNN. In: *The IEEE International Conference on Computer Vision (ICCV)*, pp. 1440–1448.
- Gaber, T., Tharwat, A., Hassanien, A.E., Snasel, V., 2016. Biometric cattle identification approach based on Weber's Local Descriptor and AdaBoost classifier. *Comput. Electron. Agric.* 122, 55–66.
- Guo, Y., Zhang, Z., He, D., Niu, J., Tan, Y.i., 2019. Detection of cow mounting behavior using region geometry and optical flow characteristics. *Comput. Electron. Agric.* 163, 104828. <https://doi.org/10.1016/j.compag.2019.05.037>.
- García, R., Aguilar, J., Toro, M., Pinto, A., Rodríguez, P., 2020. A systematic literature review on the use of machine learning in precision livestock farming. *Comput. Electron. Agric.* 179, 105826. <https://doi.org/10.1016/j.compag.2020.105826>.
- Hertem, T., Viazzi, S., Steensels, M., Maltz, E., Antler, A., Alchanatis, V., Schlageter-Tello, A.A., Lokhorst, K., Romanini, E.C.B., Bahr, C., Berckmans, D., Halachmi, I., 2014. Automatic lameness detection based on consecutive 3D-video recordings. *Biosyst. Eng.* 119, 108–116.
- Hansen, M.F., Smith, M.L., Smith, L.N., Salter, M.G., Baxter, E.M., Farish, M., Grieve, B., 2018. Towards on-farm pig face recognition using convolutional neural networks. *Comput. Ind.* 98, 145–152.
- Huang, W., Zhu, W., Ma, C., Guo, Y., Chen, C., 2018. Identification of group-housed pigs based on Gabor and Local Binary Pattern features. *Biosyst. Eng.* 166, 90–100.
- Hao, S., Zhou, Y., Guo, Y., 2020. A Brief Survey on Semantic Segmentation with Deep Learning. *Neurocomputing* 406, 302–321.
- Huang, W., Zhu, W., Ma, C., Guo, Y., 2020. Weber texture local descriptor for identification of group-housed pigs. *Sensors* 20, 4649.
- Hu, H., Dai, B., Shen, W., Wei, X., Sun, J., Li, R., Zhang, Y., 2020. Cow identification based on fusion of deep parts features. *Biosyst. Eng.* 192, 245–256.
- Jiang, B.o., Wu, Q., Yin, X., Wu, D., Song, H., He, D., 2019. FLYOLOv3 deep learning for key parts of dairy cow body detection. *Comput. Electron. Agric.* 166, 104982. <https://doi.org/10.1016/j.compag.2019.104982>.
- Jiang, B.o., Yin, X., Song, H., 2020. Single-stream long-term optical flow convolution network for action recognition of lameness dairy cow. *Comput. Electron. Agric.* 175, 105536. <https://doi.org/10.1016/j.compag.2020.105536>.
- Jia, N., Kootstra, G., Koerkamp, P.G., Shi, Z., Du, S., 2021. Segmentation of body parts of cows in RGB-depth images based on template matching. *Comput. Electron. Agric.* 180, 105897. <https://doi.org/10.1016/j.compag.2020.105897>.
- Kashiha, M., Bahr, C., Haredasht, S.A., Ott, S., Moons, C.P.H., Niewold, T.A., Ödberg, F.O., Berckmans, D., 2013. The automatic monitoring of pigs water use by cameras. *Comput. Electron. Agric.* 90, 164–169.
- Khoramshahi, E., Hietaja, J., Valros, A., Yun, J., Pastell, M., 2014. Real-time recognition of sows in video: a supervised approach. *Inform. Process. Agric.* 1 (1), 73–81.
- Kumar, S., Pandey, A., Sai Ram Satwik, K., Kumar, S., Singh, S.K., Singh, A.K., Mohan, A., 2018. Deep learning framework for recognition of cattle using muzzle point image pattern. *Measurement* 116, 1–17.
- Lu, Y., He, X., Wen, Y., Wang, P.S.P., 2014. A new cow identification system based on iris analysis and recognition. *IJBM* 6 (1), 18. <https://doi.org/10.1504/IJBM.2014.059639>.
- Lu, M., Xiong, Y., Li, K., Liu, L., Yan, L.i., Ding, Y., Lin, X., Yang, X., Shen, M., 2016. An automatic splitting method for the adhesive piglets' gray scale image based on the ellipse shape feature. *Comput. Electron. Agric.* 120, 53–62.
- Lee, J., Jin, L., Park, D., Chung, Y., 2016. Automatic recognition of aggressive behavior in pigs using a kinetic depth sensor. *Sensors* 16, 631.
- Lao, F., Brown-Brandl, T., Stinn, J.P., Liu, K., Teng, G., Xin, H., 2016. Automatic recognition of lactating sow behaviors through depth image processing. *Comput. Electron. Agric.* 125, 56–62.
- Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C.Y., Berg, A.C., 2016. SSD: Single Shot MultiBox Detector. In: *Proceedings of the European Conference on Computer Vision, Amsterdam, The Netherlands, 8-16 October 2016*; Springer, Berlin/Heidelberg, Germany, 2016, pp. 21–37.

- Li, W., Ji, Z., Wang, L., Sun, C., Yang, X., 2017a. Automatic individual identification of Holstein dairy cows using tailhead images. *Comput. Electron. Agric.* 142, 622–631.
- Lahrman, H.P., Hansen, C.F., DEath, R.B., Busch, M.E., Nielsen, J.P., Forkman, B., 2018. Early intervention with enrichment can prevent tail biting outbreaks in weaner pigs. *Livestock Sci.* 214, 272–277.
- Larsen, M.L.V., Andersen, H.-M.-L., Pedersen, L.J., 2019. Changes in activity and object manipulation before tail damage in finisher pigs as an early detector of tail biting. *Animal* 13 (5), 1037–1044.
- Leonard, S.M., Xin, H., Brown-Brandt, T.M., Ramirez, B.C., 2019. Development and application of an image acquisition system for characterizing sow behaviors in farrowing stalls. *Comput. Electron. Agric.* 163, 104866. <https://doi.org/10.1016/j.compag.2019.104866>.
- Li, B.o., Liu, L., Shen, M., Sun, Y., Lu, M., 2019a. Group-housed pig detection in video surveillance of overhead views using multi-feature template matching. *Biosyst. Eng.* 181, 28–39.
- Li, D., Chen, Y., Zhang, K., Li, Z., 2019b. Mounting behaviour recognition for pigs based on deep learning. *Sensors* 19, 4924.
- Li, X., Cai, C., Zhang, R., Ju, L., He, J., 2019c. Deep cascaded convolutional models for cattle pose estimation. *Comput. Electron. Agric.* 164, 104885. <https://doi.org/10.1016/j.compag.2019.104885>.
- Liu, H.e., Reibman, A.R., Boerman, J.P., 2020a. Video analytic system for detecting cow structure. *Comput. Electron. Agric.* 178, 105761. <https://doi.org/10.1016/j.compag.2020.105761>.
- Liu, D., Oczak, M., Maschat, K., Baumgartner, J., Pletzer, B., He, D., Norton, T., 2020b. A computer vision-based method for spatial-temporal action recognition of tail-biting behavior in group-housed pigs. *Biosyst. Eng.* 195, 27–41.
- Li, D., Zhang, K., Li, Z., Chen, Y., 2020. A spatiotemporal convolutional network for multi-behavior recognition of pigs. *Sensors* 20, 2381.
- Larsen, M.L.V., Wang, M., Norton, T., 2021. Information technologies for welfare monitoring in pigs and their relation to welfare quality®. *Sustainability* 13, 692.
- Matthews, S.G., Miller, A.L., Clapp, J., Plötz, T., Kyriazakis, I., 2016. Early detection of health and welfare compromises through automated detection of behavioural changes in pigs. *Veterinary J.* 217, 43–51.
- Morris, M.J., Kaneko, K., Walker, S.L., Jones, D.N., Routly, J.E., Smith, R.F., Dobson, H., 2011. Influence of lameness on follicular growth, ovulation, reproductive hormone concentrations and estrus behavior in dairy cows. *Theriogenology* 76 (4), 658–668.
- Muns, R., Manzanilla, E.G., Sol, C., Manteca, X., Gasa, J., 2013. Piglet behavior as a measure of vitality and its influence on piglet survival and growth during lactation. *J. Animal Sci.* 91 (4), 1838–1843.
- Marcon, M., Brossard, L., Quiniou, N., 2015. Precision feeding based on individual daily body weight of group-housed pigs with an automatic feeder developed to allow for restricting feed allowance. *Prec. Livest. Farm.* 15, 592–601.
- Marsot, M., Mei, J., Shan, X., Ye, L., Feng, P., Yan, X., Li, C., Zhao, Y., 2020. An adaptive pig face recognition approach using Convolutional Neural Networks. *Comput. Electron. Agric.* 173, 105386. <https://doi.org/10.1016/j.compag.2020.105386>.
- Nasirahmadi, A., Richter, U., Hensel, O., Edwards, S., Sturm, B., 2015. Using machine vision for investigation of changes in pig group lying patterns. *Comput. Electron. Agric.* 119, 184–190.
- Nasirahmadi, A., Hensel, O., Edwards, S.A., Sturm, B., 2016. Automatic detection of mounting behaviours among pigs using image analysis. *Comput. Electron. Agric.* 124, 295–302.
- Norton, T., Chen, C., Larsen, M.L.V., Berckmans, D., 2019. Review: Precision livestock farming: building 'digital representations' to bring the animals closer to the farmer. *Animal* 13 (12), 3009–3017.
- Nasirahmadi, A., Sturm, B., Olsson, A.-C., Jeppsson, K.-H., Müller, S., Edwards, S., Hensel, O., 2019a. Automatic scoring of lateral and sternal lying posture in grouped pigs using image processing and Support Vector Machine. *Comput. Electron. Agric.* 156, 475–481.
- Nasirahmadi, A., Sturm, B., Edwards, S., Jeppsson, K.H., Olsson, A.C., Müller, S., Hensel, O., 2019b. Deep learning and machine vision approaches for posture detection of individual pigs. *Sensors* 19, 3738.
- Oczak, M., Viazzi, S., Ismayilova, G., Sonoda, L.T., Roulston, N., Fels, M., Bahr, C., Hartung, J., Guarino, M., Berckmans, D., Vranken, E., 2014. Classification of aggressive behaviour in pigs by activity index and multilayer feed forward neural network. *Biosyst. Eng.* 119, 89–97.
- Ouared, K., Zidane, K., Aggad, H., Niar, A., 2015. Impact of clinical lameness on the milk yield of dairy cows. *J. Animal Veterinary Adv.* 14, 10–12.
- Okura, F., Ikuma, S., Makihara, Y., Muramatsu, D., Nakada, K., Yagi, Y., 2019. RGB-D video-based individual identification of dairy cows using gait and texture analyses. *Comput. Electron. Agric.* 165, 104944. <https://doi.org/10.1016/j.compag.2019.104944>.
- Poursaberi, A., Bahr, C., Pluk, A., Van Nuffel, A., Berckmans, D., 2010. Real-time automatic lameness detection based on back posture extraction in dairy cattle: shape analysis of cow with image processing techniques. *Comput. Electron. Agric.* 74 (1), 110–119.
- Porto, S.M.C., Arcidiacono, C., Anguza, U., Cascone, G., 2013. A computer vision-based system for the automatic detection of lying behaviour of dairy cows in free-stall barns. *Biosyst. Eng.* 115 (2), 184–194.
- Porto, S.M.C., Arcidiacono, C., Anguza, U., Cascone, G., 2015. The automatic detection of dairy cow feeding and standing behaviours in free-stall barns by a computer vision-based system. *Biosyst. Eng.* 133, 46–55.
- Psota, E.T., Mittek, M., Pérez, L.C., Schmidt, T., Mote, B., 2019. Multi-pig part detection and association with a fully-convolutional network. *Sensors* 19, 852.
- Qiao, Y., Truman, M., Sukkariéh, S., 2019a. Cattle segmentation and contour extraction based on Mask R-CNN for precision livestock farming. *Comput. Electron. Agric.* 165, 104958. <https://doi.org/10.1016/j.compag.2019.104958>.
- Qiao, Y., Su, D., Kong, H.e., Sukkariéh, S., Lomax, S., Clark, C., 2019b. Individual cattle identification using a deep learning based framework. *IFAC-PapersOnLine* 52 (30), 318–323.
- Redmon, J., Divvala, S., Girshick, R., Farhadi, A., 2016. You only look once: unified, real-time object detection. In: *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 779–788.
- Rieker, M., Klein, A., Adrion, F., Hoffmann, C., Gallmann, E., 2020. Automatically detecting pig position and posture by 2D camera imaging and deep learning. *Comput. Electron. Agric.* 174, 105391. <https://doi.org/10.1016/j.compag.2020.105391>.
- Prashanth, C.R., Sudarshan, T.S.B., 2020. Deep learning methods for multi-species animal re-identification and tracking-a survey. *Comput. Sci. Rev.* 38, 100289.
- Song, X., Leroy, T., Vranken, E., Maertens, W., Sonck, B., Berckmans, D., 2008. Automatic detection of lameness in dairy cattle—vision-based trackway analysis in cow's locomotion. *Comput. Electron. Agric.* 64 (1), 39–44.
- Srivastava, N., Mansimov, E., Salakhutdinov, R., 2015. Unsupervised learning of video representations using LSTMs. *International Conference on Machine Learning*.
- Sun, Y., Huo, P., Wang, Y., Cui, Z., Li, Y., Dai, B., Li, R., Zhang, Y., 2019. Automatic monitoring system for individual dairy cows based on a deep learning framework that provides identification via body parts and estimation of body condition score. *J. Dairy Sci.* 102, 10140–10151.
- Turner, S.P., Farnworth, M.J., White, I.M.S., Brotherstone, S., Mendl, M., Knap, P., Penny, P., Lawrence, A.B., 2006. The accumulation of skin lesions and their use as a predictor of individual aggressiveness in pigs. *Appl. Animal Behav. Sci.* 96 (3–4), 245–259.
- Tu, G.J., Karstoft, H., Pedersen, L.J., Jørgensen, E., 2013. Foreground detection using loopy belief propagation. *Biosyst. Eng.* 116 (1), 88–96.
- Teixeira, D.L., Boyle, L.A., 2014. A comparison of the impact of behaviours performed by entire male and female pigs prior to slaughter on skin lesion scores of the carcass. *Livestock Sci.* 170, 142–149.
- Tian, M., Guo, H., Chen, H., Wang, Q., Long, C., Ma, Y., 2019. Automated pig counting using deep learning. *Comput. Electron. Agric.* 163, 104840. <https://doi.org/10.1016/j.compag.2019.05.049>.
- Tu, S., Liu, H., Li, J., Huang, J., Li, B., Pang, J., & Xue, Y. (2020). Instance segmentation based on mask scoring R-CNN for group-housed pigs. In: *2020 International Conference on Computer Engineering and Application (ICCEA)*, pp. 458–462.
- Viazzi, S., Bahr, C., Schlageter-Tello, A., Van Hertem, T., Romanini, C.E.B., Pluk, A., Halachmi, I., Lokhorst, C., Berckmans, D., 2013. Analysis of individual classification of lameness using automatic measurement of back posture in dairy cattle. *J. Dairy Sci.* 96 (1), 257–266.
- Viazzi, S., Ismayilova, G., Oczak, M., Sonoda, L.T., Fels, M., Guarino, M., Vranken, E., Hartung, J., Bahr, C., Berckmans, D., 2014a. Image feature extraction for classification of aggressive interactions among pigs. *Comput. Electron. Agric.* 104, 57–62.
- Viazzi, S., Bahr, C., Van Hertem, T., Schlageter-Tello, A., Romanini, C.E.B., Halachmi, I., Lokhorst, C., Berckmans, D., 2014b. Comparison of a three-dimensional and two-dimensional camera system for automated measurement of back posture in dairy cows. *Comput. Electron. Agric.* 100, 139–147.
- Vila, R.M., Tummaruk, P., 2016. Management strategies in farrowing house to improve piglet pre-weaning survival and growth. *Thai J. Veterinary Med.* 46 (463), 347–354.
- Wu, D., Wu, Q., Yin, X., Jiang, B.o., Wang, H., He, D., Song, H., 2020. Lameness detection of dairy cows based on the YOLOv3 deep learning algorithm and a relative step size characteristic vector. *Biosyst. Eng.* 189, 150–163.
- Xu, B., Wang, W., Falzon, G., Kwan, P., Guo, L., Chen, G., Tait, A., Schneider, D., 2020. Automated cattle counting using Mask R-CNN in quadcopter vision system. *Comput. Electron. Agric.* 171, 105300. <https://doi.org/10.1016/j.compag.2020.105300>.
- Yazdanbakhsh, O., Zhou, Y.u., Dick, S., 2017. An intelligent system for livestock disease surveillance. *Inf. Sci.* 378, 26–47.
- Yang, A., Huang, H., Zheng, C., Zhu, X., Yang, X., Chen, P., Xue, Y., 2018a. High-accuracy image segmentation for lactating sows using a fully convolutional network. *Biosyst. Eng.* 176, 36–47.
- Yang, Q., Xiao, D., Lin, S., 2018b. Feeding behavior recognition for group-housed pigs with the Faster R-CNN. *Comput. Electron. Agric.* 155, 453–460.
- Yang, A., Huang, H., Zhu, X., Yang, X., Chen, P., Li, S., Xue, Y., 2018c. Automatic recognition of sow nursing behaviour using deep learning-based segmentation and spatial and temporal features. *Biosyst. Eng.* 175, 133–145.
- Yang, A., Huang, H., Yang, X., Li, S., Chen, C., Gan, H., Xue, Y., 2019. Automated video analysis of sow nursing behavior based on fully convolutional network and oriented optical flow. *Comput. Electron. Agric.* 167, 105048. <https://doi.org/10.1016/j.compag.2019.105048>.
- Yang, Q., Xiao, D., 2020. A review of video-based pig behavior recognition. *Appl. Animal Behav. Sci.* 233, 105146. <https://doi.org/10.1016/j.applanim.2020.105146>.
- Yang, A., Huang, H., Zheng, B., Li, S., Gan, H., Chen, C., Yang, X., Xue, Y., 2020. An automatic recognition framework for sow daily behaviours based on motion and image analyses. *Biosyst. Eng.* 192, 56–71.
- Yin, X., Wu, D., Shang, Y., Jiang, B.o., Song, H., 2020. Using an EfficientNet-LSTM for the recognition of single Cow's motion behaviours in a complicated environment. *Comput. Electron. Agric.* 177, 105707. <https://doi.org/10.1016/j.compag.2020.105707>.
- Zhu, W.-X., Guo, Y.-Z., Jiao, P.-P., Ma, C.-H., Chen, C., 2017. Recognition and drinking behaviour analysis of individual pigs based on machine vision. *Livestock Sci.* 205, 129–136.
- Zhao, K., Bewley, J.M., He, D., Jin, X., 2018. Automatic lameness detection in dairy cattle based on leg swing analysis with an image processing technique. *Comput. Electron. Agric.* 148, 226–236.

- Zheng, C., Zhu, X., Yang, X., Wang, L., Tu, S., Xue, Y., 2018. Automatic recognition of lactating sow postures from depth images by deep learning detector. *Comput. Electron. Agric.* 147, 51–63.
- Zhao, K., Jin, X., Ji, J., Wang, J., Ma, H., Zhu, X., 2019. Individual identification of Holstein dairy cows based on detecting and matching feature points in body images. *Biosyst. Eng.* 181, 128–139.
- Zhang, L., Gray, H., Ye, X., Collins, L., Allinson, N., 2019a. Automatic individual pig detection and tracking in pig farms. *Sensors* 19, 1188.
- Zhang, Y., Cai, J., Xiao, D., Li, Z., Xiong, B., 2019b. Real-time sow behavior detection based on deep learning. *Comput. Electron. Agric.* 163, 104884. <https://doi.org/10.1016/j.compag.2019.104884>.
- Zhu, X., Chen, C., Zheng, B., Yang, X., Gan, H., Zheng, C., Yang, A., Mao, L., Xue, Y., 2020. Automatic recognition of lactating sow postures by refined two-stream RGB-D faster R-CNN. *Biosyst. Eng.* 189, 116–132.
- Zheng, C., Yang, X., Zhu, X., Chen, C., Wang, L., Tu, S., Yang, A., Xue, Y., 2020. Automatic posture change analysis of lactating sows by action localisation and tube optimization from untrimmed depth videos. *Biosyst. Eng.* 194, 227–250.
- Zhang, K., Li, D., Huang, J., Chen, Y., 2020. Automated video behavior recognition of pigs using two-stream convolutional networks. *Sensors* 20, 1085.