DETECTION OF BEHAVIOUR AND POSTURE OF SHEEP BASED ON YOLOv3 / 基于YOLOv3的绵羊行为姿态检测

Xuefeng Deng*[1)], Xiaoli Yan[2)], Yiming Hou[1)], Hui Wu[1)], Chenru Feng[1)], Lingyu Chen[1)], Maoxing Bi[1)], Yi Shao[1)] [1]

[1)] College of Information Science and Engineering, Shanxi Agricultural University, Taigu / China

[2)] Taiyuan Jinshan Middle School, Taiyuan / China

Tel: 13546632115; E-mail: dxf@sxau.edu.cn

DOI: https://doi.org/10.35633/inmateh-64-45

Keywords: *YOLOv3, Posture recognition, Object detection, Sheep posture*

ABSTRACT

The behaviour and posture of animals are closely related to their physiological conditions. To some extent, we can judge their physiological activity by their behaviour and posture. Sheep's behaviour change significantly during illness or parturition, these behaviours are composed of simple postures, such as standing, eating, lying down. This paper takes the YOLOv3 algorithm as the core technology. It extracts features of sheep's behaviour and posture through constructing the deep network structure, and uses the pyramid feature fusion and multi-scale prediction to detect the behaviour and posture of sheep. The experimental results show that the training model can effectively detect the three behaviours and postures of sheep: standing, eating, and lying down. The mean average precision is 92.47%. This experiment can be used as a basic technology to judge the physiological activities of sheep. It can be applied to the intelligence of animal husbandry, and has a broad application prospect.

摘要

动物的行为姿态与其生理状况息息相关，通过动物的行为姿态可以在一定程度上判断其生理活动。绵羊在出现疾病或分娩时其行为会有显著的变化，这些行为由站立、进食和躺卧等简单姿态的组合而成，本文以YOLOv3算法作为核心技术，通过构建深层网络结构，提取绵羊行为姿态特征，采用金字塔特征融合和多尺度的预测对绵羊行为姿态进行检测。经实验表明，训练的模型可以有效检测出绵羊的站立、进食、躺卧三种行为姿态，平均精度均值达到了92.47%。本实验可以作为判断绵羊生理活动的一种基础技术，可以应用到畜牧业智能化中，具有广阔的应用前景。

INTRODUCTION

The behaviour of animals is closely related to their physiological conditions. The behaviour detection of animals can judge their physiological activities to a certain extent *(He et al., 2016).* It is the core content of accurate animal husbandry. For example, feeding, excretion and drinking behaviour of animals are an important basis for judging animal health. When animals excrete, feed and drink too frequently, this means that they may have diseases such as viral diarrhoea and infectious gastroenteritis *(Zhu et al., 2010)*. When the animal often appears chewing without food, lying down, walking back and forth, constantly looking back at the abdomen and other behaviours, it means that the animal is during pregnancy. Accurate detection and analysis of animal postures is a basic process of animal behaviour detection. Animal behaviours show different combinations of postures. Sheep is a common breeding animal. In the process of feeding, posture recognition of sheep can make a relatively objective evaluation of its growth, health and pregnancy state, and can timely take effective measures such as prevention and treatment, human intervention and other effective measures, so as to reduce the loss to the minimum.

With the development of artificial intelligence technology, intelligent methods can be used to detect and classify animal postures. Because traditional animal husbandry relies on the continuous observation of the breeders and their own experience to judge illness or parturition time of animals and so on, this will cause heavy labour burden, low work efficiency, and also increase the transmission rate of zoonosis, so it is very necessary to use intelligent means to conduct behaviour detection.

[1] *Xuefeng Deng*, assoc.prof, PhD; Xiaoli Yan, Lec, M.Eng.; Yiming Hou, undergraduate; Hui Wu, undergraduate; Chenru Feng, undergraduate; Lingyu Chen, undergraduate; Maoxing Bi, undergraduate; Yi Shao, undergraduate.*

At present, there is a variety of algorithms widely used in animal posture detection.

Xue et al., (2018), solved the difficulties in posture recognition of suckling sows caused by the changes of light and thermal light lines in the environment with day and night alternation and the attachment between sows and piglets in the free pen scene. Taking the deep video image as the data source, they proposed an improved Faster R-CNN posture recognition algorithm for suckling sows. That is, through end-to-end training network, at the same time, the detection of lactating sows and the classification of five postures, such as standing, sitting, prone lying down, abdominal lying down, side lying down and so on, are realized. Although the speed of the algorithm is fast, the accuracy is not high.

To solve the problem of the low degree of automatic detection of pig crossing behaviour at present, *Li et al., (2019),* proposed a crawling behaviour of pigs recognition algorithm based on Mask R-CNN. By obtaining the top view image of the pig, using LabelImg to make the data set label, and introducing the transfer learning method to train the ResNet-FPN network, the pig segmentation results were obtained and the mask pixel area of each sample is extracted, but the algorithm is slow in image processing and takes a long time.

Ye et al., (2019), proposed a test method based on Faster-RCNN. When broilers are in stun state, make use of Convolution neural network to extract broilers' features. The speed of this algorithm can meet production needs, but its training speed is low and its occupancy space is big.

Zhuang et al., (2020), proposed a convolutional neural network algorithm for recognizing the oestrus behaviour of large white sows in view of the characteristics such as the erect ears of large white sows during the oestrus test. By collecting the ear images of large white sows in oestrus test and non-oestrus test, the training set samples (80%) and validation set samples (20%) were divided for later training. The algorithm has a high recognition rate, but its speed is still low and has the disadvantages of spending too long time. In recent years, intelligent methods have been introduced in feeding sheep to analysis and research sheep's behaviour, so it can promote production process intelligentization. For example, *Fogarty et al. (2020)* proposed a sheep behaviour classification method based on machine learning. It can identify sheep's behaviour such as grazing, lying down, standing, walking and so on by machine learning. The combination of these postures can identify sheep's many behaviours.

Sun et al. (2019) research sheep parturition behaviour and get Faster-RCNN model based on Soft-NMS algorithm and VGG16 feature extraction network which can detect sheep's parturition behaviour well. *Salama et al. (2019)* proposed a sheep identification method based on deep learning and Bayes optimization. It uses Bayes optimization algorithm to set convolutional neural network parameter automatically, so it can identify sheep's face in order to achieve the goal of identifying sheep.

Ma et al. (2020) researched sheep identification location and proposed a Faster-FCNN neural network model based on Soft-NMS algorithm, which can real time monitor and locate sheep in complex raise conditions.

In sheep research above artificial intelligence algorithms there are problems such as computational efficiency and running space, during animal identification and monitoring.

This paper adopts YOLOv3 to identify sheep's postures. It uses Labellmg tool to annotate sheep's postures such as standing, feeding, lying down. Through training, it generates YOLOv3 network which can identify relevant postures. This algorithm has the characteristics of fast detection speed and high accuracy. It is close to real-time detection and provides a technical solution of artificial intelligence method applied to sheep feeding.

MATERIALS AND METHODS

Image acquisition of behaviour and posture of sheep

Collection of experimental data on a sheep farm in Lianbai Village, Hejin City, Shanxi Province, which covers an area of about 40m*25m and raises a variety of sheep such as Australian white sheep, Suffolk sheep, Dorper sheep, Small-tailed Han sheep and Hybrid sheep. Considering the influence of different illumination on the establishment of the model, we take photos in the morning and afternoon respectively to ensure that the number of pictures under different lighting conditions is the same.

Since April 3, 2021, three hours of pictures have been taken in the morning and afternoon, for five consecutive days, and 1500 images have been collected as data sets. The images collected include three behaviours and postures of sheep: eating, standing and lying down, as shown in Figure 1.

Fig. 1 - Three Behaviours and Postures of Sheep

Sheep postures image data preprocessing

Data preprocessing mainly includes the following processes: data filtering, data enhancement and data annotation. The data preprocessing process is shown in Figure 2.

Data filtering mainly refers to the uniform selection of pictures of sheep with different behaviours and postures in different time periods, different illumination intensity and different shooting directions. 1000 pictures were selected in this experiment. Data enhancement mainly involves random flipping, random rotation, random clipping, and random scaling of the image. In this experiment, some pictures are randomly symmetrically flipped so as to enhance the state information of sheep in different positions. Data annotation is to classify and label the original image. The corresponding labels of different postures such as eating, standing and lying down are marked respectively, preparing for further data processing.

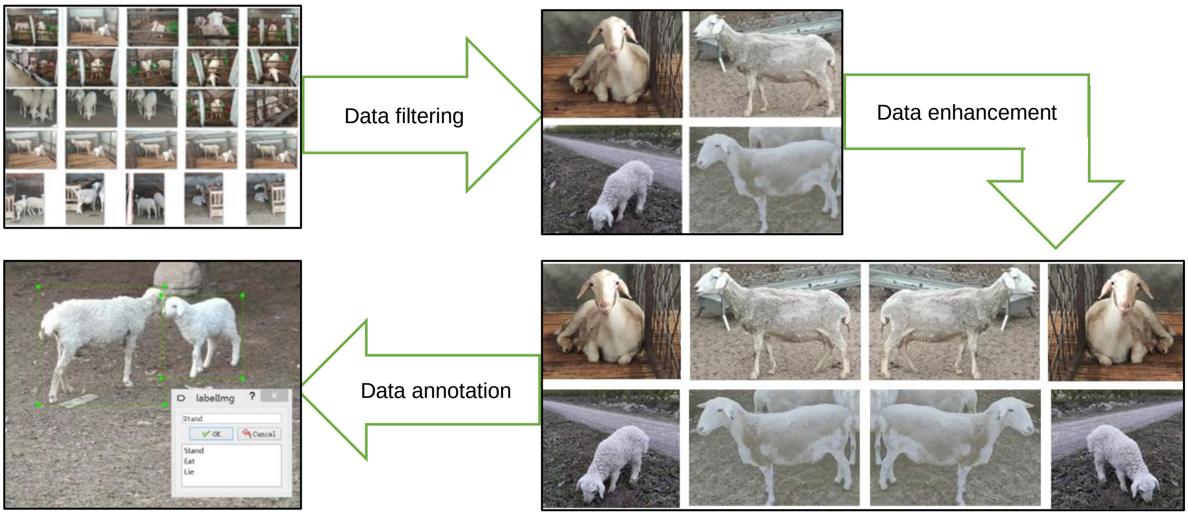

Fig. 2 - Data Preprocessing Process of Sheep Posture Images

Introduction of behaviour and posture recognition methods for sheep

This paper uses the YOLOv3 algorithm and Darknet-53 network model to detect behaviour and posture of sheep. The flow of sheep behaviour and posture detection method is shown in Figure 3.

In the data input part, the pre marked data is used as input. The Labellmg image annotation tool was mainly used to label the standing, eating and lying behaviour and posture of sheep. Then save the label information marked on the picture to an XML file accordingly. In the part of constructing feature extraction network, a reasonable feature extraction network is selected to meet the needs of detection. In this paper, D is used as the network structure for posture feature extraction. Feature fusion and posture prediction mainly use YOLOv3 algorithm to build the relevant network structure and complete the later processing. Finally, after selecting the optimal prediction result, input the posture detection result of sheep.

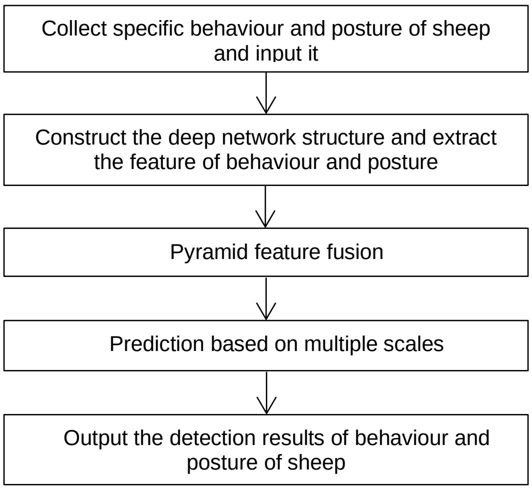

Fig. 3 - Detection Process of Behaviour and Posture of Sheep

RESEARCH METHOD OF SHEEP POSTURE DETECTION

Introduction of YOLOv3 algorithm

As we all know, a more accurate and faster object detection algorithm is the primary principle of our algorithm selection. YOLO series algorithm is a popular algorithm in the field of object detection, YOLOv3 is a masterpiece of integrated SSD (multi-scale prediction), FCN (full volume machine), FPN (feature pyramid), DenseNet network (feature channel concat). It is a typical representative of a single-stage object detection algorithm. Compared with two-stage algorithms such as Fast RCNN, the YOLO series algorithm abandons the step of candidate region extraction, and only uses a one-level network to complete object detection and location, which can complete the object detection task at near real-time speed*(Sun et al., 2021)*.

YOLOv3 is an object detection model based on Darknet-53 proposed by Joseph Redmon et al. The essence of object detection is recognition and regression. YOLOv3 combines the Darknet-53 network model with the feature pyramid model as the backbone network. At the same time, it also introduces the more advanced Resnet residual network, which is helpful to solve the problem of gradient disappearance and the explosion of a deep network *(Pan et al., 2021).* The method of feature pyramid is used in the object detection of theYOLOv3 algorithm. The feature maps of different depths are detected respectively, so that the feature maps of the current layer will sample and use the feature maps of the future layer. The small-size feature map can detect large-size objects, and the large-size feature map can detect the role of small-size objects. Compared with YOLOv1 and YOLOv2, YOLOv3 improves the accuracy and speed of object detection, completes the adjustment of network structure, uses multi-scale features for object detection, and replaces softmax with logistic in object classification *(Tan et al., 2021)*.

Principle of posture discrimination of sheep base on YOLOv3

The basic structure of object detection by YOLOv3 is shown in Figure 4. The basic structure includes input layer, convolution layer, feature fusion layer and output layer, with 107 layers in total, of which 0-74 layers are convolution layer; 75-105 layers are responsible for object detection, classification and regression. In the input layer, we choose to input the image of 416×416×3. YOLOv3 network is with Darknet-53 network model as convolutional layers to sample under 32 times, so input layers should be three-channel RGB image of n × n, and n should be a multiple of 32. The Convolutional layer is from 0 to 74th layer, and it consists of residual modules and convolutional modules sampling between residual modules. The convolution modules sample under 32 times, 16 times and 8 times use steps of two, which effectively reduces the loss of low-level features. Adding a residual module can ensure that the network structure can continue to converge in a very deep situation to a certain extent. At the same time, the amount of calculation is reduced. The feature fusion layer is 75-105, and 416×416×3 images as input, object detection, mosaic feature information, fusion 13 ×13, 26 × 26, 52 ×52. Each scale interacts with local features through a convolution kernel to complete pyramid feature fusion. The output layer is mainly responsible for the classification and position regression of the three scale feature maps output by the feature fusion layer. These three feature maps will be transferred into the logistic layer respectively.

The reason why the softmax layer is not used is that the softmax layer can only generate one classification for each frame, but in our object detection task, there may be multiple target objects overlapping. In this way, there will be multiple classifications, so logistics is used in YOLOv3 to generate the output of the model. There are 13 ×13 ×3 + 26 ×26 ×3 + 52 ×52 ×3 predictions in YOLOv3 *(Zhang et al., 2021)*.

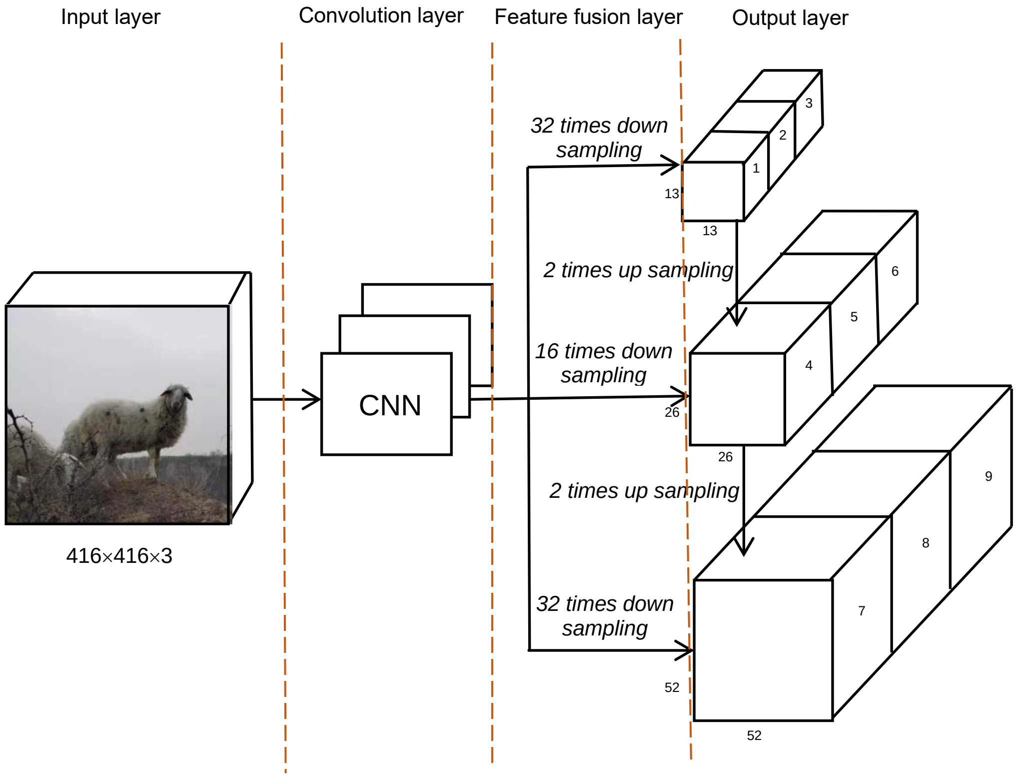

Fig. 4 - Basic Structure of Detection by YOLOv3

Darknet-53 network structure

Based on the idea of Resnet, compared with Darknet-19 network, the accuracy and speed of Darknet-53 network processing a large number of picture data have been greatly improved. Because the network structure contains 53 convolution layers, it is named Darknet-53, which is integrated from the convolution layers with better performance selected from various mainstream network structures. In ImageNet image classification, the processing speed of Darknet-53 is 78 images per second, which is much faster than Resnet with the same precision, twice the efficiency of Resnet-152, and the effect is better than Darknet-19. Darknet-53 adds a residual module into the network, which has a convolution core of 32 filters and 5 groups of residual units. Each unit of these 5 groups of residual units is composed of a separate convolution layer and a group of repeated convolution layers. The repeated convolution layers are repeated once, twice, eight times, eight times, four times respectively, then perform a series of convolution layers of 3 × 3 and 1 × 1. The reason why the residual module is added is that it can solve the gradient problem of deep network. Because the deeper the network is, the phenomenon of gradient disappearing will become more and more obvious, and the training effect of the network will not be very good. The residual neural network is to solve the problem of gradient disappearance under the condition of deepening the network. The residual structure does not pass convolution, mapping directly from the previous feature layer to the following feature layer (jump connection) is helpful for training and feature extraction, which is easy to optimize *(Sun et al., 2021)*.

FORMATION PROCESS OF CLASSIFICATION OF SHEEP BEHAVIOUR AND POSTURE

Feature extraction of sheep image

The purpose of this paper is to detect the object of three kinds of designated behaviour and posture of sheep. In order to recognize accurately and quickly, we adopt the YOLOv3 algorithm as the core technology. The core idea of YOLOv3 object detection is to output the detected object information at one time, including category and location.

It only needs to find out which grid the centre of the object is in, and it does not need many ASK like Fast-RCNN. It can be summarized as the following steps: First, input sheep image, YOLOv3 will set it to 416 × 416 images, and add a blackness bar to prevent distortion; second, the image will be divided into three kinds of grid images, which are 13 × 13, 26 × 26 and 52 × 52. Among them, 13 × 13 is mainly used for large object detection, 26 × 26 is used for medium object detection, 52 × 52 is used for small object detection, and 13 × 13 grids are selected for sheep image recognition; third, if the centre of the bounding box corresponding to ground truth in the training set just falls into a grid cell of the input image, then the grid cell is used to predict the object; fourth, because each grid cell will predict a fixed number of bounding boxes, the largest bounding box is used to predict the object. The steps of YOLOv3 object detection are shown in Figure 5 *(Cai et al., 2020)*.

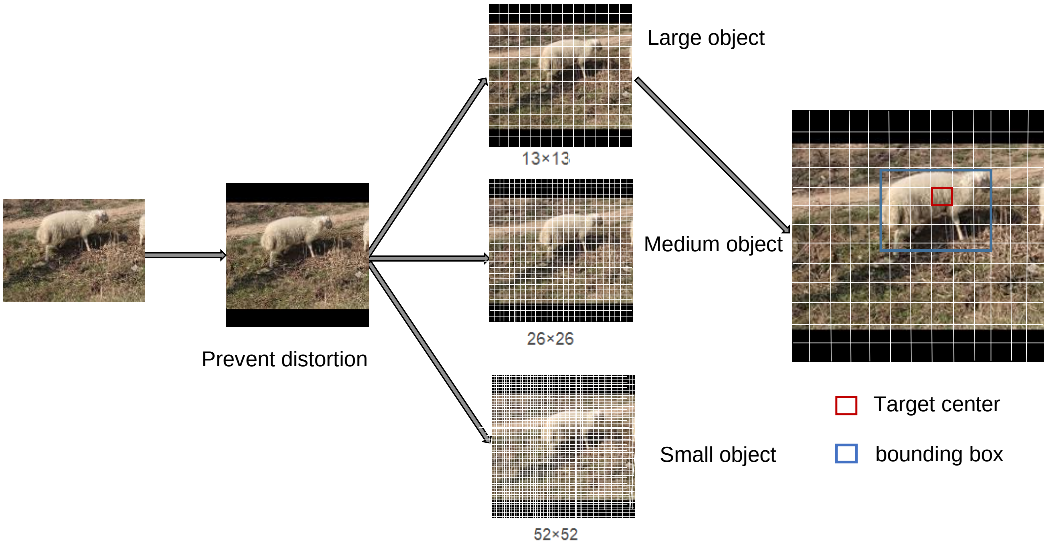

Fig.5 - Detection Steps of Sheep Based on YOLOv3

Network recognition of sheep posture features and calculation method

YOLOv3 algorithm is based on intersection and union ratio as one of the important indexes of the evaluation model, the union ratio refers to the union ratio between the generated candidate box and the original marked box. It can be used to evaluate the object detection algorithm and judge whether it works well. It is the standard performance measure of object class segmentation problem, and its expression is

$$R_{IOU} = \frac{I\ (X)}{U\ (X)} \tag{1}$$

Where $I(x)=\Sigma_{v\in V}XvYv$, $U(X)=\Sigma_{v\in V}(Xv+Yv-XvYv)$

$V=\{1, 2, \ldots, N\}$ is the pixel set of all the images of behaviour and posture of sheep in the training set; X is a network output for the pixel probability on the set V; Y is the actual pixel probability situation on the set V.

The main idea of calculating the bounding box in YOLOv3 is to divide the input picture of behaviour and posture of sheep into uniform equivalent cells, and then calculate the coordinate position of the cell where the centre is located, so as to realize the prediction of the bounding box. YOLOv3 predicts four coordinate values for each bounding box (t_x, t_y, t_w, t_h). For the predicted grid cell, according to the offset of the upper left corner of the image (c_x, c_y) and the width and height of the bounding box (P_x, P_y) to predict the bounding box in the following way, the calculation formula is as follows:

$$b_x = \delta(t_x) + c_x \tag{2}$$

$$b_y = \delta(t_y) + c_y \tag{3}$$

$$b_w = P_w\ e^{t_w} \tag{4}$$

$$b_h = P_h\ e^{t_h} \tag{5}$$

Among them, t_x, t_y, t_w, t_h is the predicted four coordinate values; t_x, t_y represents the offset of the predicted coordinates; t_w, t_h represents the scaling of the dimensions respectively;

P_x, P_y represents the width and height of the prediction bounding box;

c_x, c_y is the upper left corner coordinate of a grid cell, which represents the width and height of the cell, that is, the coordinate offset;

$\delta(t_x)$, $\delta(t_y)$ represents the cell coordinates of the centre point in the cell in which it is located; b_x, b_y, b_w, b_h represents the coordinate position and size of the bounding box relative to the input image, that is, the predicted output coordinates *(Qu et al., 2021).*

YOLOv3 algorithm uses K-means clustering to determine the size of the annotation box, and continues the method of YOLOv2. Three kinds of prior boxes are set for each down sampling scale, and a total of 9 kinds of size prior boxes are clustered. When calculating the dimension box, the coordinates x and y of the centre point of all the dimension boxes are set to 0, so that all the dimension boxes are placed in the same position, to a certain extent, it can be more effective to calculate the similarity between the annotation boxes, the clustering distance measurement function is as follows:

$$\mathrm{D} = 1 - R_{IOU} \tag{6}$$

The last level of convolution of the YOLOv3 algorithm is 1 × 1, so the calculation formula of convolution kernel size is:

$$l = [\mathrm{B}\times（5 + \mathrm{C}）\times 1\times 1] \tag{7}$$

Where B is predicting the number of bounding boxes for the network; C is the number of categories; 5 is the four coordinate values and one confidence level, for this experiment to detect behaviour and posture of sheep, C will be set to 3, B will be set to 3.

The confidence level of the YOLOv3 algorithm can be used to evaluate the reliability of the test image. The intersection and union ratio of the probability of an object contained in a grid cell and the object detection is the confidence level of the YOLOv3 algorithm. The calculation formula is as follows:

$$F_{con} = Pr(Obj)R_{IOU} \tag{8}$$

Where the value of the $Pr(Obj)$ is determined by the probability of whether the prediction box contains the object to be detected. When the sheep is detected as one of the three postures of eating, standing and lying down, it will be set to 1, otherwise it will be set to 0; R_{IOU} represents the union ratio, as can be seen above.

Realization process of behaviour and posture of sheep

The process of detection model of behaviour and posture of sheep based on YOLOv3 is as follows, as shown in Figure 6:

I. Construction of sheep behaviour and posture data set.

The data set of behaviour and posture of sheep was obtained by human shooting. From the 1500 pictures taken, 1000 pictures with good shooting effect were selected as the input data set of this experiment, including three postures of sheep standing, eating and lying down, and the number of pictures of these three behaviours and postures was equal. Then the images will be resized and normalized. Because there are 5 downsampling steps of 2 in the Darknet-53, the feature map will be reduced by 32 times. So we need to resize the image to a multiple of 32. In this experiment, we resize the image of sheep data set to 416×416. In order to speed up the convergence of the training network, the image is normalized from 0-255 pixels to 0-1 pixels without changing the image information. Then LabelImg is used to label the image to accurately mark the specific position of the corresponding category of objects in the image *(Zhao et al., 2021)*.

II. Model construction and training detection model

(1) The datasets are divided, dividing the training set and test set at 9:1. 900 pictures are randomly extracted as the training set and 100 images as the test set.

(2) The YOLOv3 network is configured, and the main YOLOv3 network parameters adjusted in this experiment include: after each iteration, the number of images sent to the network is batch=16. Each batch is divided into the number of subdivision corresponding numbers, the number of calculations set to subdivisions = 16. The width of the input image is width=416, and the height of the input image is height=416. The number of objects identified by the network is set to classes=3. The filters in the last convolutional layer before each layer is set to filters=24.

(3) Download the pre-trained weight darknet53.conv.74 before starting the training. The training process includes: higher resolution training for backbone networks, backbone and detection networks combined, reading training data and preprocessing, defining loss functions, model training.

After training the backbone network, we add the corresponding convolutional layer based on this network, realize the architecture of image feature pyramid (FPN), and then output three different eigenvalues of different image resolutions, and finally build a network detection system that can subsample the pictures by 8 times, 16 times and 32 times.

This model takes the three outputs route1, route2 and route3 of the backbone network darknet-53 as input, and the final output predict is the prediction result of three different dimensions.

III. To verify the accuracy of the model.

This experiment uses loss curves and Region Avg IOU curve to evaluate model training effects. At last, mean average precision (mAP) is calculated from the test set data to verify the availability of the model.

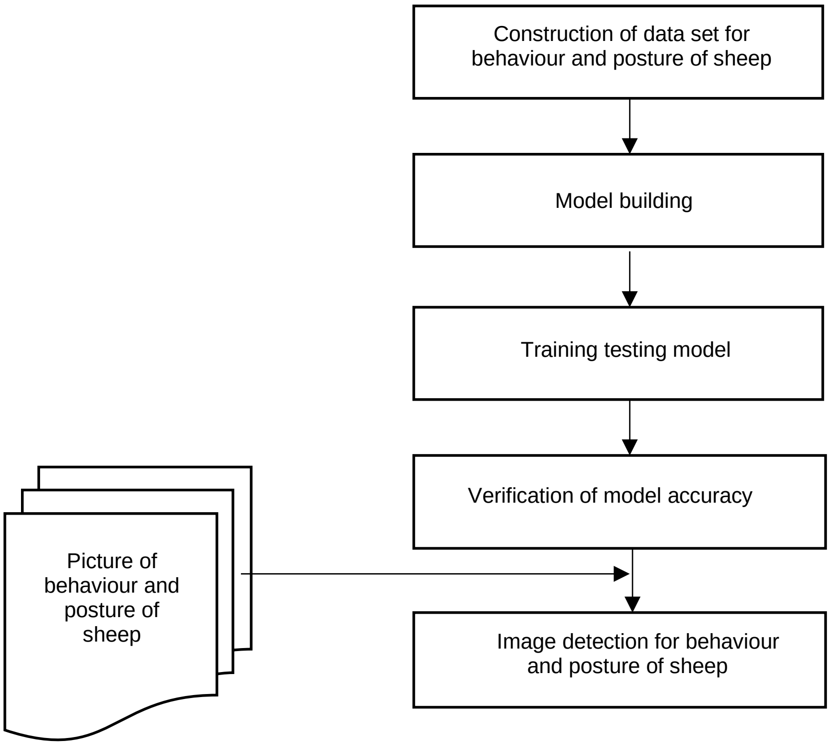

Fig. 6 - Flow Chart of Detection Model of Behaviour and Posture of Sheep

CLASSIFICATION OF SHEEP BEHAVIOUR AND POSTURE

The verification effect of this experiment is shown in Figure 7. The following images in Figure 7 are selected in the test set by random. When we put these pictures in our training model of sheep's behaviour and posture, it can accurately identify sheep's three postures such as eating, standing and lying down, according to the characteristics of the input images. It can be seen from the effect in the figure that this method can distinguish the relevant posture of a single sheep. If there are multiple target objects in a scene, it can also achieve good results.

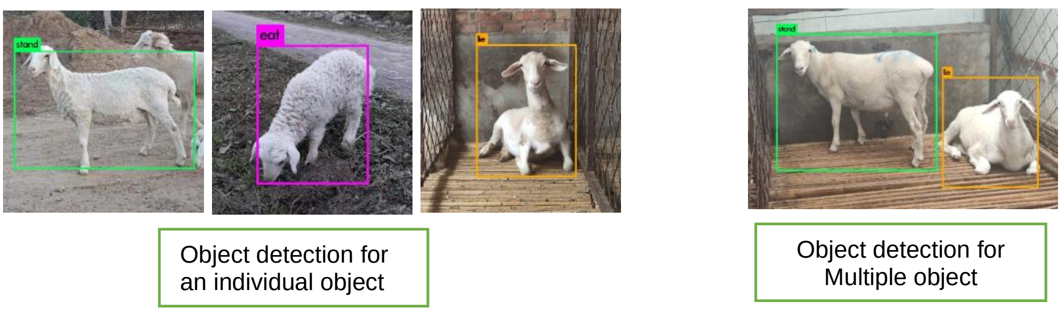

Fig. 7 – Classification of Sheep Behaviour and Posture

By analysing the output data in the training process, the loss curves and Region Avg IOU curves can be drawn. The Loss curves is shown in Figure 8. In this experiment, when the training times exceed 800 times, the loss value has been lower than the set threshold, which proves that the training can be stopped and the results can be tested. The region AVG IOU curve is shown in Figure 9. It can be seen from the curve that when the training times reach 600, the IOU value has approached 1, indicating that the coincidence degree between the predicted rectangular box and the target in the training model detection results has met the requirements.

By analysing the test set and using the detector valid command of darknet network, the test results of three categories to be detected in this experiment can be obtained in the darknet/results directory. Through the comparison between the test results and the annotation in the test set, the average precision of various postures is calculated respectively (standing posture is 84.38%, eating posture is 95.74%, lying posture is 97.29%). So the Mean Average Precision (mAP) is 92.47%. It can meet the detection requirements in accuracy.

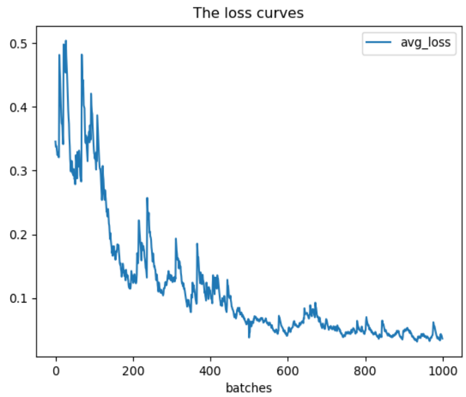

Fig. 8 – The Loss Curves

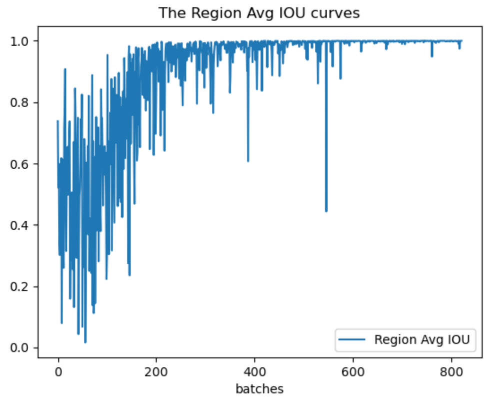

Fig. 9 – The Region Avg IOU Curves

CONCLUSIONS AND PROSPECTS

In view of the situation that behaviour and posture of sheep will change significantly when they suffer from illness or parturition, this paper uses the YOLOv3 algorithm as the core to realize the detection of behaviour and posture of sheep, which can effectively and accurately identify their three behaviours: standing, eating and lying down. By predicting the behaviour and posture of animals in advance, we can effectively judge the physiological activities of animals, which is helpful for breeders to stop the abnormal behaviour of sheep in advance. It also saves human resources and enhances the production and economic benefits of agriculture.

The detection results in this paper have achieved good results for individual posture detection. However, when multiple individuals are mixed, if there is no mutual occlusion, the algorithm can accurately detect posture of multiple target individuals. However, when different individuals block each other or some sheep are blocked by other obstacles, the detection results of blocked individuals are not accurate. This is also a problem to be studied in the future. Yolov3 is a real-time target detection tool. After verification, the method in this paper can be added to the real-time video processing system for the application of real-time monitoring scene, which provides a practical and feasible scheme for the application in related fields.

ACKNOWLEDGEMENT

This research, titled 'DETECTION OF BEHAVIOUR AND POSTURE OF SHEEP BASED ON YOLOv3', was funded by the research project of Shanxi Province excellent doctor work Award Fund (Research on model detection technology of Agricultural Internet of things, No: SXYBKY201721), Scientific Research Start of Doctor of Shanxi Agricultural University (Research on Detection Technology of Agricultural Internet of Things Model, 2017YJ30). The authors are grateful and honoured to receive this support.

REFERENCES

[1] Cai F.H., Zhang Y.X., Huang J., (2020), Detection Algorithm of Bridge Surface Crack Based on YOLOv3 and Attention Mechanism. *Pattern Recognition and Artificial Intelligence*, 10, 33, 926-933.

[2] Fogarty E.S., Swain D.L., Cronin G.M., Moraes L.E., Trotter M., (2020), Behaviour classification of extensively grazed sheep using machine learning. *Computers and Electronics in Agriculture*, 169: 105175.
[3] He D.J., Liu D., Zhao K.X., (2016), Review of Perceiving Animal Information and Behaviour in Precision Livestock Farming. *Journal of Agricultural Machinery*, 05, 47, 231-244.
[4] Li D., Zhang K.F., Li X.J., Chen Y.F., Li Z.B., Pu D., (2019), Mounting Behaviour Recognition for Pigs Based on Mask R-CNN. *Journal of Agricultural Machinery*, S1, 50, 261-266+275.
[5] Ma C., Sun X., Yao C., et al. (2020), Research on Sheep Recognition Algorithm Based on Deep Learning in Animal Husbandry. *Journal of Physics: Conference Series.* IOP Publishing, 1651(1): 012129.
[6] Pan Y.H., Wei J.S., Zeng L.P., (2021), Farmland bird detection algorithm based on YOLOv3. *Progress in Laser and Optoelectronics*, 1-16.
[7] Qu J.S., Zhu Z.Q., Wan Q.B., (2021), Face Key Point Detection Based on YOLOv3. *Yangtze River Information Communication*, 01, 34, 55-58.
[8] Salama A., Hassanien A.E., Fahmy A., (2019), Sheep identification using a hybrid deep learning and Bayesian optimization approach. *IEEE Access*, 7: 31681-31687.
[9] Sun Q.Y., Zhang Z.D., (2021), Crowd Density Estimation Based on YOLOv3 Enhanced Model Fusion. *Application of Computer Systems*, 04, 30, 271-276.
[10] Sun S., Qin J., Xue H., (2019), Sheep delivery scene detection based on faster-RCNN. *The Second International Conference on Image, Video Processing and Artificial Intelligence.*
[11] Tan F., Mu P.A., Ma Z.X., (2021), Multi-object Tracking Algorithm Based on YOLOv3 Detection and Feature Point Matching. *Journal of Computing*, 02, 42, 157-162.
[12] Xue Y.J., Zhu X.M., Zheng C., Mao L., Yang A.Q., Tu S.Q., Huang N., Yang X.F., Chen P.F., Zhang N.F., (2018), Lactating Sow Postures Recognition From Depth Image of Videos Based on Improved Faster R-CNN. *Transactions of the Chinese Society of Agricultural Engineering*, 09, 34, 189-196.
[13] Ye C.W., Kang R., Qi C., Liu C., Zhao Y., Chen K.J., (2019), Detection method of broiler strike based on Faster-RCNN. *Journal of Agricultural Machinery*, 12, 50, 255-259.
[14] Zhang X.C.，Zhao J.S., Wang S.Z., Zhang M, Cheng C., (2021), Binocular Vision Detection and Positioning Method for Ships Based on YOLOv3 Algorithm. *Journal of Shanghai Maritime University*, 01, 42, 26-32.
[15] Zhao Y.Y., Zhu J., Xie Y.K., Li W.L., Guo Y.K., (2021), Improved Yolo-v3 Video Image Flame Real-time Detection Algorithm. *Geomatics and Information Science of Wuhan University*, 03, 46, 326-334.
[16] Zhuang Y.R., Yu J.H., Teng G.H., Cao M.B., (2020), Study on the Detection of Estrus Behaviour in Big White Pig Based on Convolutional Neural Network. *Journal of Agricultural Machinery*, S1, 51, 364-370.
[17] Zhu W.X., Pu X.F., Li X.C., Lu C.F., (2010), Automatic identification system of suspected sick pigs based on behavioural monitoring. *Journal of Agricultural Engineering*, 01, 26, 188-192.