DETECTION OF EATING BEHAVIOUR IN PIGS BASED ON MODIFIED YOLOX / 基于改进的 YOLOX 猪只饮食行为检测

Yanwen LI, Juxia LI *), Lei DUAN, Tengxiao NA, Pengpeng ZHANG, Qingyu ZHI
College of Information Science and Engineering, Shanxi Agricultural University/China
Tel: +86-15803446486; E-mail: lijxsn@126.com
DOI: https://doi.org/10.35633/inmateh-71-03

Keywords: *deep learning, behavior detection, YOLOX, attentional mechanism, pig*

ABSTRACT
Due to the complex environment of pig farms and the diversity of pig behaviors, the existing methods based on deep learning cannot meet the requirements of high accuracy and real-time detection of pig eating behavior. In this paper, a SE-YOLOX model for detecting pig diet and drinking behaviors was designed. In this model, a Squeeze-and-Excitation (SE) attention module is introduced between the neck layer and the prediction layer of YOLOX, and the input feature map is compressed into a vector through global average pooling operation, and then mapped to a smaller vector through a fully connected layer. A sigmoid function is also used to compress each element in this vector to between 0 and 1 and multiply it with the original input feature map to get the weighted feature map. Through SE attention mechanism, the model can learn the importance of each channel adaptively, thus improving the detection accuracy. The experimental results show that the mean Average Prediction (mAP) of the SE-YOLOX model is 88.03%, which is higher than 13.11% of the original YOLOX model. SE-YOLOX can ensure real-time performance, it also can effectively improve the accuracy of pig diet and drinking water behavior detection.

摘要
由于养猪场环境的复杂和猪行为的多样性，现有的基于深度学习的方法难以满足猪食用行为检测的高精度和实时性要求，本文提出了一种 SE-YOLOX 猪只饮食饮水检测模型。该模型在 YOLOX 的 Neck 层和 Prediction 层之间引入 Squeeze-and-Excitation (SE)注意力模块，通过全局平均池化操作将输入特征图压缩成一个向量，然后通过一个全连接层将其映射到一个较小的向量；同时使用一个 sigmoid 函数将这个向量中的每个元素压缩到 0 到 1 之间，并将其与原始输入特征图相乘，得到加权后的特征图。通过 SE 注意力机制，该模型可以自适应地学习到每个通道的重要特征，从而提高检测精度。实验结果表明，SE-YOLOX 模型的平均检测精度(mAP)为 88.03%，高于原 YOLOX 模型 13.11%。SE-YOLOX 在保证实时性的同时，能有效提高猪饮食饮水行为检测的准确性。

INTRODUCTION
The eating behavior of pigs can often reflect the health and emotional problems of pigs and is an important basis for judging the health of pigs. Through monitoring, the dietary behavior rules of pigs can be extracted, and feed amount, water consumption, and other resources can be allocated rationally according to different times and seasons. Artificial monitoring of pigs' dietary behavior requires a large amount of manpower and is difficult to achieve long-term and sustained observation. Meanwhile, errors in judgment will inevitably occur. Automatic monitoring of pig-eating behavior has become a hot topic in recent years. It not only improves the level of automatic pig breeding but also has great significance to master the health status of pigs and improve their growth welfare (*Chen et al., 2022; Statham et al., 2020; Shen et al., 2014*).

At present, the commonly used methods to detect pig eating behavior include radio frequency identification technology and machine vision technology. RFID technology is mainly used to identify the eating behavior of pigs by wearing electronic ear tag sensors. In addition, RFID receivers are installed next to drinking water and eating utensils. When pigs are close to drinking and eating utensils, the receiver reads the corresponding information, thereby recording pig eating behavior and pig health problems (*Maselyne et al., 2014; Li et al., 2021; Wang, 2021*). Although RFID technology has high accuracy and can complete real-time monitoring of pigs' dietary behavior, each pig needs to wear ear tags, which leads to increased costs. At the same time, the sensor is damaged or lost due to abnormal behavior such as ear biting.

Nowadays, Machine vision technology based on deep learning can not only accurately identify pig behavior, but also has the advantages of convenience and low cost, and has been widely used in the current pig-raising process (*Qin et al., 2021)*. Yang Qiumei achieved pig detection through image segmentation, and the accuracy rate of pig drinking behavior recognition reached 92.11% (*Yang et al., 2018*). Wu Shihai used convolutional neural networks to extract the depth features of pig images, and the accuracy of pig behavior recognition in sequence images reached 94% (*Wu et al., 2020*). The above target detection model mainly adopts the setting of a priori frame, which cannot well solve the influence of factors such as the size of the target in the image and the mutual occlusion of the target on the detection result during model detection.

In order to ensure the real-time detection of the pig dietary behavior detection model, the detection accuracy is improved as much as possible. In this paper, a targeted method for pig diet behavior detection based on the SE-YOLOX model is proposed. This model ensures the accuracy of real-time detection and can be ported to various devices, which can be applied to various aquaculture farms.

MATERIALS AND METHODS

Data Acquisition and Preprocessing

The experimental data in this paper were collected from the pig breeding base in Fenxi County, Linfen City, Shanxi Province. In order to obtain the diet images of pigs in different environments, four adult pigs were selected and placed in a captive environment of 3.0m×3.0m×2.0m from July 2, 2020. 3h videos were filmed in the morning, afternoon, and night for 7 consecutive days. The total video length is 63h. Allied Vision Technologies' Manta G-282C camera was used. In order to clearly check the eating and drinking situation of pigs during the day, a stationary camera is one meter high from the ground and forms a 35° angle to the wall which is mounted on the wall to the right of the pig feed trough. In order to prevent unsupervised shooting at night and reduce the difficulty of shooting, another camera with a height of 3 meters and forms a 45° angle to the wall was placed on the roof of the piggery for night shooting.

KMPlayer software was used for video interception of the collected data, and JPG images were extracted according to 1F/s for the captured video, and then the obtained images were screened. In order to enrich the training data of the model, the selected samples included samples of different time periods, different lighting intensities and different shooting angles, and finally 1615 images were selected as the training and test sample data. The sample size was uniformly adjusted to 640 *640 pixels to improve the training speed of the model.

After selecting appropriate samples, LabelImg image labeling tool was used to label the eating behavior of pigs, as shown in Fig. 1.

Fig. 1 - Data annotation

SE-YOLOX algorithm

YOLOX is a new high-performance detector based on YOLOv3, which improved the Anchor free mechanism, the Decoupled Head, and the label distribution strategy SimOTA. It achieved significant performance improvements while still maintaining the efficient reasoning characteristics of the YOLO series. Its network structure is shown in Fig. 2.

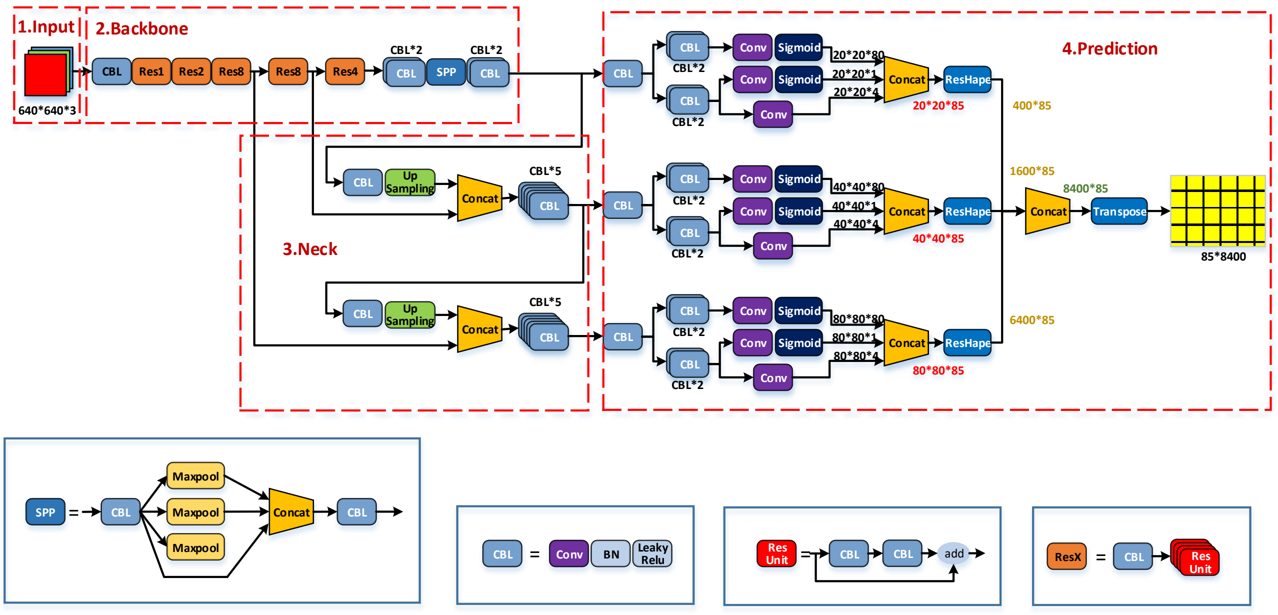


Fig. 2 - YOLOX network structure

The backbone network of YOLOX network structure is Darknet53, in which neck layer and prediction layer are updated and strengthened. YOLOX adopts Mosaic and MixUp data enhancement methods on the input side. For neck layer, YOLOX still adopts the structure of Feature Parymid Network (FPN) for fusion, and transmits and fuses the feature information of high level through up-sampling from top to bottom, so as to obtain the feature graph for prediction. YOLOX achieves a better balance in terms of speed and accuracy than any other model of YOLO series.

In order to obtain more detailed information about the target of attention from different sources and suppress other useless information, the Squeeze and Excitation (SE) Networks attention mechanism was introduced. SE attention mechanism is a channel type of attention mechanism. It's main content is squeeze and excitation. Through automatic learning, another new neural network is used to obtain the importance of each channel of the feature map and then use this importance to assign a weight value to each feature, so that the neural network can focus on certain feature channels. Promote the feature graph channels that are useful to the current task, and suppress the feature channels that are not useful to the current task. The main innovation of the network is that the model can automatically learn the importance of different channel features by focusing on the relationship between channels. The SE attention mechanism mainly includes operations through squeezes and excitation. The squeeze operation uses global averaging pooling to encode the entire spatial features on the channel as local features.

The calculation method is shown in (1) :

$$z_c = \frac{1}{H \times W} \sum_{i=1}^{H} \sum_{j=1}^{w} u_c(i, j) \tag{1}$$

In this formula, the second two-dimensional matrix in the convolution 3d matrix represents the result of the squeeze operation, and the subscript represents the number of channels. After the squeeze operation gets the channel information, use two full connection layers to form the gate mechanism and activate it with Sigmoid function. The calculation method is shown in (2) :

$$s = F_{ex}(z, W) = \sigma(g(z, W)) = \sigma(W_2 \delta(W_1 z)) \tag{2}$$

Where:

δ is ReLU activation function, σ is Sigmoid function, W_1 and W_2 are the weights used by the two fully connected layers to reduce and increase dimensions, respectively equal to $\frac{C}{r} \times C$ and $\frac{C}{r} \times C$, and r is the scaling parameter to limit model complexity and improve model capability. s represents the weight set of feature graphs obtained through the fully connected layer and nonlinear layer. Finally, the weight of the output is assigned to the original feature. The calculation formula is shown in (3):

$$\tilde{x}_c = s_c \times u_c \tag{3}$$

where, $\tilde{x}_c$ is the feature graph of the feature channel of x, s_c is the weight, and u_c is a two-dimensional matrix. As shown in Figure 3:

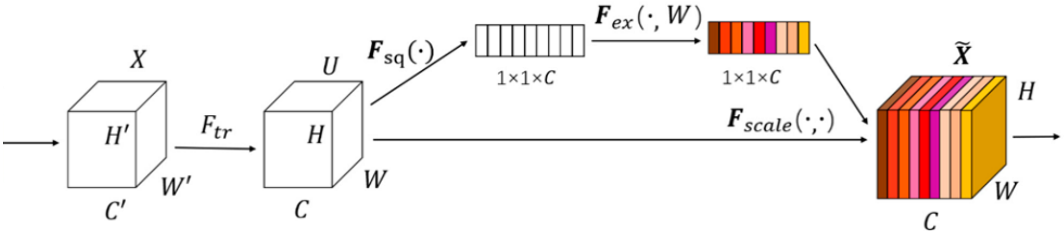


Fig. 3 - The structure of SE attention mechanism

This example inserts SE attention mechanism in the YOLOX. The network structure name is defined as SE-YOLOX. The SE-YOLOX network structure is shown in Fig.4.

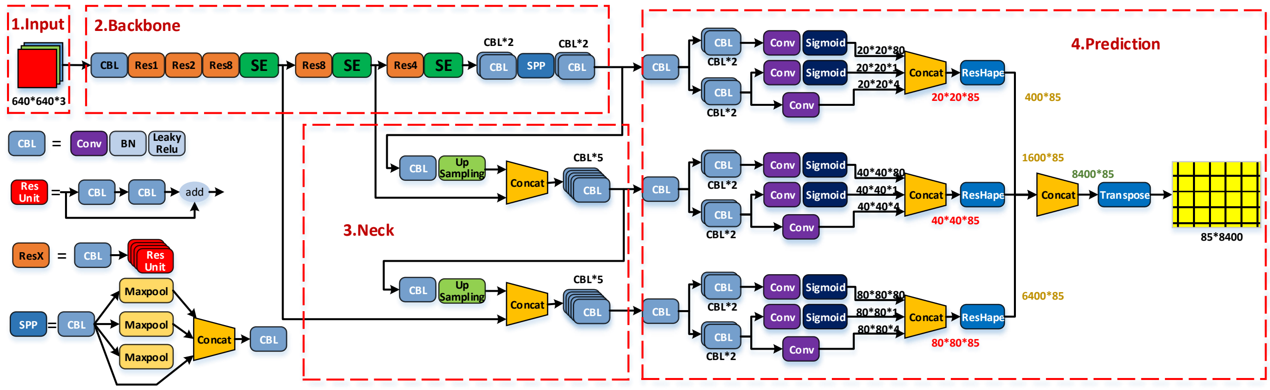


Fig. 4 - SE-YOLOX network structure

Evaluation indexes of the model

In this paper, commonly used evaluation indexes in target detection field include IoU, Precision, Recall, P-R curve, AP value and mAP value.

When evaluating the target detection model, IoU is used to quantify the degree of fitting, that is, to judge the quality of detection from the degree of fitting between the prediction frame and the real frame. The calculation formula is shown in (4) :

$$IoU = \frac{S_A \cap S_B}{S_A \cup S_B} \tag{4}$$

where S_A represents the collection of pixels in the prediction frame area, S_B represents the collection of pixels in the real frame area.

Precision and Recall are used to evaluate the quality of an information retrieval system. Precision stands for detection prediction, representing the proportion of targets detected by the model to be real target objects. Recall stands for detection recall rate, representing the proportion of all real targets effectively detected by the model. Precision and Recall formulas are shown in (5) and (6) :

$$Precision = \frac{TP}{TP + FP} \tag{5}$$

$$Recall = \frac{TP}{TP + FN} \tag{6}$$

TP (True Positive) refers to the number of samples whose IoU is greater than the threshold and whose category label is consistent with the actual label. FP (False Positive) represents the number of samples predicted by the model as individual boxes of pigs and whose category label is inconsistent with the actual label. FN (False Negative) indicates the number of samples for which no individual pig is detected.

The Average Precision (AP) of the model detection target is the area under the P-R curve, and the average precision can be used to evaluate the overall performance of the model for target detection and classification. The mean of average accuracy represents the average of the average accuracy of all classes. Compared with the mean of average accuracy, the mean of average accuracy can more accurately reflect the overall performance of the model to detect various targets. AP is the measure of the area formed by P–R curve and coordinate axis and is given as Formula (7):

$$AP = \int_0^1 p(r)dr \tag{7}$$

where p is the abbreviation of Precision and r is the abbreviation of Recall.

The mAP value can be calculated by

$$mAP = \frac{\sum_{i=1}^{N} AP_i}{N} \tag{8}$$

where N is the number of images in the test set. The mAP represents the mean average precision of the model. It can evaluate the overall detection effect of the model on the test set.

Considering that the selection of different IoU thresholds directly affects TP and FP values and thus causes the fluctuation of mAP index values, in this paper, three IoU thresholds of 0.5, 0.75, and 0.5:0.05:0.95 (where 0.05 represents the growth step) were denoted as $mAP_{0.5}$, $mAP_{0.75}$, and $mAP_{0.5:0.95}$ respectively to measure the model detection under different conditions. At the same time, considering the accurate detection performance of the model, the median and maximum of mAP under 10 thresholds with IoU of 0.5~0.95 are $mAP_{0.5:0.95\text{-medium}}$ and $mAP_{0.5:0.95\text{-large}}$.

Model Parameters

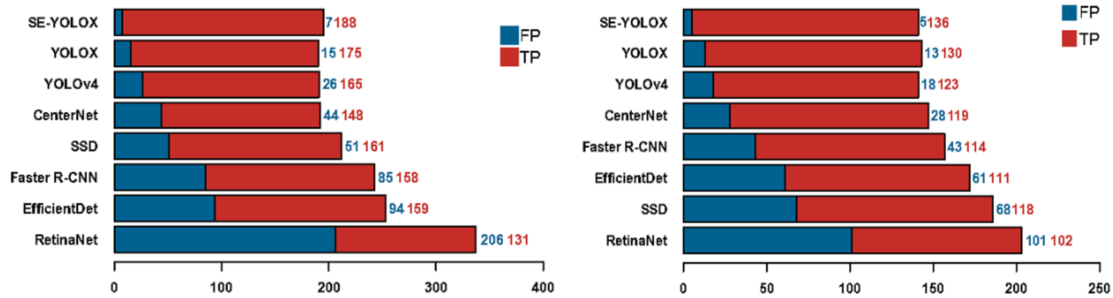
In this experiment, PyTorch deep learning framework was adopted, the operating system was Windows 10, and THE GPU was RTX2080Ti. The data set of eating and drinking images involved in this paper includes 1615 images. The collected 1615 image data sets were divided into training sets and test sets according to the ratio of 8:2, including 884 training sets and 221 test sets.

In this experiment, eight target detection models of SE-YOLOX, YOLOX (*Ge et al., 2021*), YOLOv4 (*Bochkovskiy et al., 2020*), RetinaNet (*Lin et al., 2017*), EfficientDet (*Tan et al., 2019*), Faster R-CNN (*Ren et al., 2017*), CenterNet (*Duan et al., 2019*) and SSD (*Liu et al., 2016*) were used to detect the dietary behavior of pigs on the same training set. Relevant parameters need to be set before the training of the model. In this paper, the epoch is set to 200, momentum is 0.9, and regularization coefficient for weight decay is 0.005. The experiment adopts a learning rate decay mechanism, and the initial learning rate is 0.0001. Once learning stagnates, the model learning rate usually decays at a rate of 2-10 times, and the attenuation times are set as 80% and 90% of the maximum number of iterations.

Results and analysis

In order to better evaluate the performance of the eight models involved in this paper, the same experimental data training set and test set were used for training and testing, and the results were compared and analyzed. The specific results are as follows.

TP and FP values were calculated for eight models to identify pig eating and drinking behavior. The higher the TP value, the better the model. The lower the FP value, the better the model. The results for the 8 models are presented in bar graphs, as shown in Figure 5.



a - Drink behavior **b - Eat behavior**

Fig. 5 - TP and FP values predicted by eight models

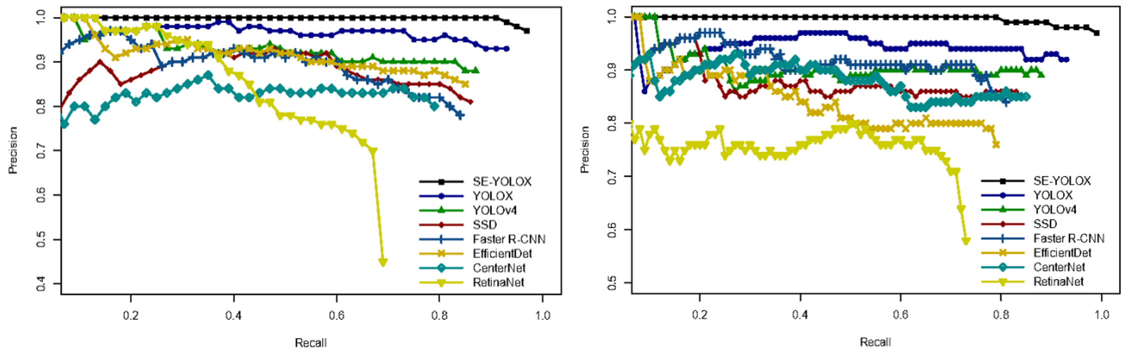
The experimental results show that:

1) As shown in Figure 5a, for the drinking behavior of pigs, the TP value of the SE-YOLOX model was significantly higher than that of the other seven models. For FP values, SE-YOLOX model detected FP values significantly lower than the other seven models. The order of FP values of the eight models was SE-YOLOX < YOLOX < YOLOv4 < CenterNet < Faster RCNN < EfficientDet < RetinaNet < SSD. The FP values of SE-YOLOX were 8 less than those of YOLOX, and 19-199 less than those of other models of the same type.

2) Figure 5b shows that for TP value, SE-YOLOX model is significantly higher than that of the other seven models. The predicted FP values of SE-YOLOX model were lower than those of the other seven models, and the order of the predicted FP values of the eight models was SE-YOLOX < YOLOX < YOLOv4 < SSD < RetinaNet < Faster RCNN < EfficientDet < CenterNet. The FP values of SE-YOLOX decreased by 8 compared with YOLOX, and significantly decreased by 13-96 compared with other models of the same type.

In conclusion, under the same experimental conditions, SE-YOLOX model is superior to other seven models in terms of TP and FP performance. Compared with the increase of TP index value, the increase of FP is much higher than that of TP, so FP is preferred in overall prediction performance to reduce the loss of detection performance caused by the large increase of FP value. It can also be seen that the improved anchor-free mode and label allocation strategy of SE-YOLOX model significantly reduced its FP value compared with other models. Therefore, SE-YOLOX model is more suitable for the detection of TP and FP values of pigs' eating and drinking behavior.

The P-R curves of the seven models for pig eating and drinking behavior recognition are shown in Figure 1. The area below the P-R curve is the AP value of the corresponding model in the eating and drinking behavior of pigs. The closer the curve is to the upper right corner, the better the model effect is. Where the black line represents the SE-YOLOX model, the dark blue line represents YOLOX model, the green line represents YOLOv4 model, the yellow line represents RetinaNet model, the orange line represents EfficientDet model, the blue line represents Faster RCNN model, the light blue line represents CenterNet model and the red line represents SSD model.



(a) drink **(b) eat**

Fig. 6 - P-R curve

The experimental results show that:

In the P-R curve prediction of pigs' drinking behavior, the P-R curve of SE-YOLOX model was always at the top, and the prediction accuracy of the model remained between 0.9 and 1.0 with the increase of Recall. The Precision of SSD model decreases rapidly when Recall is between 0.5-0.6, and the Precision of RetinaNet model decreases rapidly when Recall is between 0.6-0.7. YOLOv4, EfficientDet, Faster RCNN and CenterNet models gradually decreased with Recall of 0.8-0.9.

In the prediction of P-R curve of pig-eating behavior, the P-R curve of SE-YOLOX model was also always above other models, and the Precision value remained 1.0 between Recall and 0.8, and remained between 0.95 and 1.0 even though it decreased slightly after that. The P-R curve of SSD model decreased the fastest, and the Precision decreased rapidly when Recall was 0.4-0.5, while the Precision of the other five models decreased gradually when Recall was 0.7-0.9.

In conclusion, compared with the other six models, the P-R curves of eating and drinking of SE-YOLOX model are always at the top with stable changes, which is better than the other six models. The results indicated that the SE-YOLOX model had better stability and better detection performance for pigs' eating and drinking behaviors.

In order to explore the detection effects of different models on the eating and drinking behaviors of pigs, this paper compared the detection effects of eight models on the eating behaviors of pigs by calculating the mAP of each model. Table 1 shows the comparison of mAP of the eight models.

Table 1

Comparison of mAP values of eight models

Model	$mAP_{0.5}$	$mAP_{0.75}$	$mAP_{0.5:0.95}$	$mAP_{0.5:0.95\text{-}medium}$	$mAP_{0.5:0.95\text{-}large}$
SE-YOLOX	99.86%	98.36%	88.03%	98.36%	99.86%
YOLOX	99.95%	90.01%	74.92%	92.10%	99.95%
YOLOv4	99.88%	81.64%	68.51%	87.85%	99.88%
CenterNet	97.33%	71.99%	61.10%	76.87%	97.33%
Faster RCNN	98.71%	76.08%	63.40%	81.49%	98.71%
EfficientDet	98.79%	73.47%	61.89%	78.81%	98.79%
RetinaNet	95.50%	59.18%	54.20%	67.53%	95.50%
SSD	80.78%	45.54%	45.89%	52.63%	80.78%

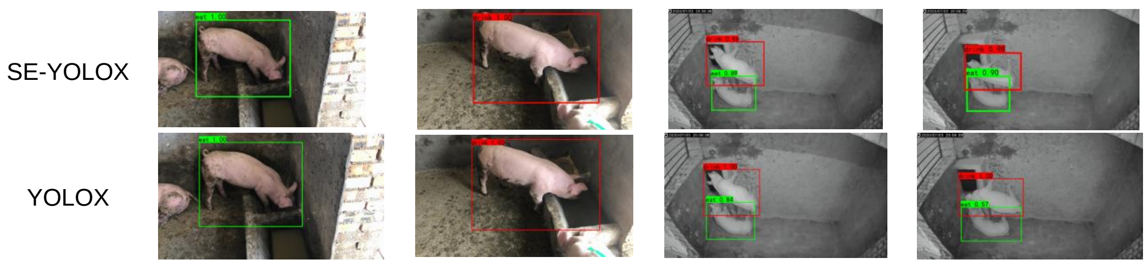
The experimental results show that:

1) Different models have different detection effects on eating and drinking behavior of pigs. As can be seen from the Table 1, the detection results of SE-YOLOX model are always higher than other models and remain in the best state. $mAP_{0.5:0.95\text{-}large}$ is 99.86%, which is only 0.09% higher than the maximum value and 1.07%-19.08% higher than other models. The $mAP_{0.5:0.95}$ of SE-YOLOX model was 13.11% higher than that of YOLOX model, 19.52% higher than that of YOLOv4 model, and higher than CenterNet and Faster RCNN, EfficientDet, RetinaNet and SSD models were 26.93%, 24.63%, 26.14%, 33.83% and 42.14%. The mAP0.5:0.95-medium of SE-YOLOX model is 98.36%, only 1.50% different from the $mAP_{0.5:0.95\text{-}large}$, while the difference between the $mAP_{0.5:0.95\text{-}medium}$ of other models and the $mAP_{0.5:0.95\text{-}large}$ is 7.85% ~ 28.85%.

2) Because the same model has different detection accuracy under different thresholds. Therefore, in the process of increasing the IoU threshold from 0.5 to 0.95, the mAP values of the eight models gradually decreased. However, the mAP value of SE-YOLOX model when IoU threshold is 0.75 is only 1.50% lower than that when IoU threshold is 0.5. YOLOX, YOLOv4, CenterNet, Faster RCNN, EfficientDet, RetinaNet and SSD models decreased 9.94%, 18.24%, 25.34%, 22.43%, 25.32%, 36.32%, 35.24%. Thus, YOLOX model is more stable in detecting the eating behavior of pigs.

In summary, it can be seen that the SE-YOLOX model has a faster convergence rate and better effect after improving the decoupling head, which not only greatly improves the detection accuracy of the experiment but also makes it more stable than other models. It can accurately predict the eating and drinking behavior of pigs at a faster speed. Therefore, SE-YOLOX model is more suitable for the detection of eating and drinking behavior of pigs.

To better evaluate the performance of the detection model, this study visualizes the results of eight detections, as shown in Figure 7. Figure 7a shows the detection results of eating behavior of pigs from a side view angle under sufficient light. Figure 7b shows the test results of drinking behavior of pigs from a side view angle under sufficient light. Figure 7c shows the detection results of the simultaneous eating and drinking behaviors of pigs from a vertical angle under the condition of insufficient light. Figure 7d shows the detection results of eating and drinking behaviors of pigs in the condition of overlooking angle adhesion under insufficient light. Part of the test results are shown in Figure 7, with the green box representing eating behavior and the red box representing drinking behavior.



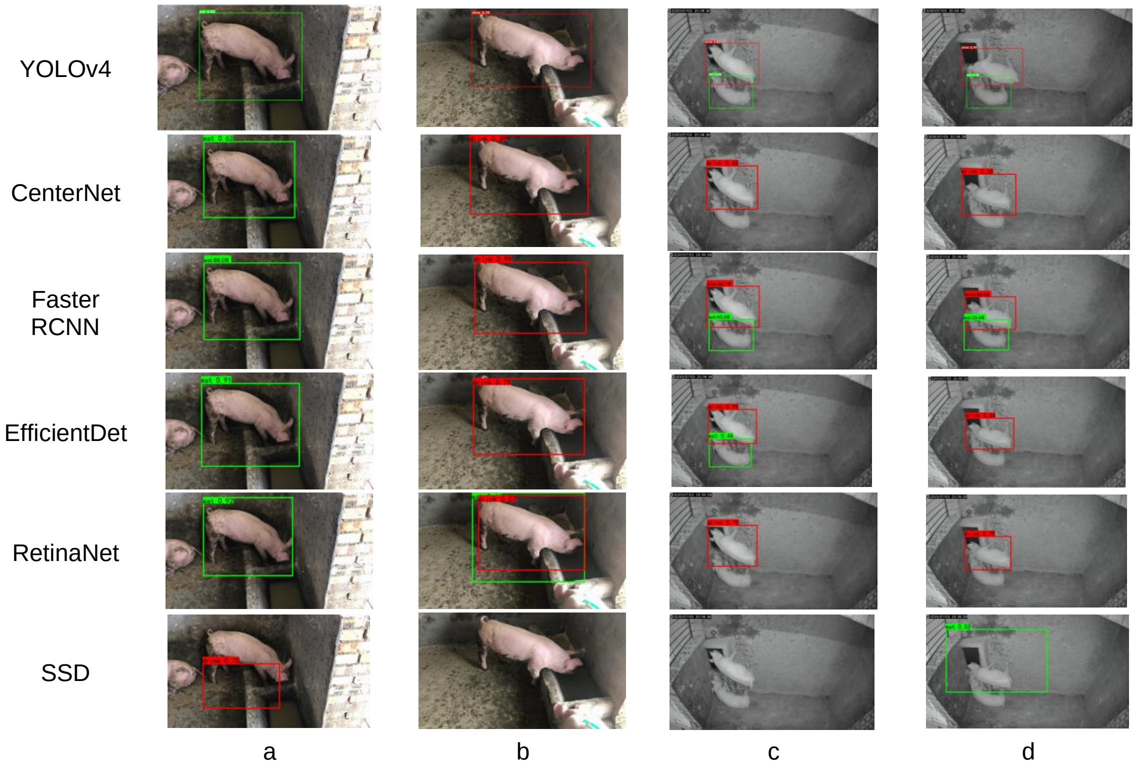


Fig. 7 - Some forecast results

Through the tests of all models, it can be seen from Fig. 7a and 7b that the eight models can detect the eating and drinking behaviors of pigs correctly under sufficient light. Among the detection results of pig eating behavior, SE-YOLOX model had the highest confidence.

Fig.7c shows that CenterNet model and RetinaNet model have errors in the detection of eating status and SSD model has errors in the detection of eating status and drinking status in the condition of insufficient light and simultaneous eating and drinking behavior. The SE-YOLOX, YOLOX, YOLOv4, Faster RCNN and EfficientDet models were all able to correctly detect drinking behavior and eating behavior, and the SE-YOLOX model had the highest confidence.

Fig.7d shows that SE-YOLOX, YOLOX, YOLOv4 and Faster RCNN models can correctly detect the eating and drinking behaviors of pigs in the sticky state under insufficient light. CenterNet, EfficientDet and RetinaNet models failed to detect the eating behavior of the pigs under the condition of adhesion, while SSD model failed to detect the eating and drinking behavior of the pigs under the condition of adhesion.

In conclusion, the SE-YOLOX model can accurately detect the eating and drinking behaviors of pigs under any circumstances, and the confidence degree is always the best compared with other models. The results showed that the SE-YOLOX model improved the input data enhancement and label allocation, and maintained the best stable state while improving the detection rate. Compared with other models, the SE-YOLOX model was more suitable for dietary detection of pigs.

CONCLUSIONS

This paper integrates the SE attention module into the YOLOX model, improves the YOLOX model, and establishes the detection model SE-YOLOX, which is used to detect the drinking and eating behavior of pigs in the pig house environment. The main conclusions are as follows:

1) Compared with the other seven deep learning detection models, the YOLOX model has better detection performance of pig drinking and eating behavior, and the model detection accuracy is improved after adding the SE attention mechanism.

2) The SE-YOLOX model outperforms the other seven models on various indicators of mAP. It indicated that the SE-YOLOX model had better stability in the detection of pigs' eating drinking behavior.

3) The SE-YOLOX model can effectively detect the eating behavior of pigs under the condition of side view, top view, adhesion state and insufficient light in real-time. While efficiently and accurately detecting the eating and drinking behavior of pigs, it does not affect the normal living order of pigs and is more easily accepted by farms. It is of great significance to the management of pig farms and the improvement of breeding welfare.

ACKNOWLEDGEMENT

This research, titled 'Detection of Eating Behavior in Pigs Based on Modified YOLOX', was funded by the Basic Research Program of Shanxi Province (202103021223141).

REFERENCES

[1] Bochkovskiy A., Wang C Y., Liao H Y M., (2020). Yolov4: Optimal speed and accuracy of object detection. *arXiv preprint arXiv*: 156,107552, India.

[2] Chen C J., Gota M., Kiho L., Zhiwu Z., Hao C., (2022). VTag: a semi-supervised pipeline for tracking pig activity with a single top-view camera. *Journal of Animal Science*, 100(6), skac147, United States.

[3] Duan K., Bai S., Xie L., et al. (2019). Centernet: Keypoint triplets for object detection. *Proceedings of the IEEE/CVF international conference on computer vision,* pp. 6569-6578, United States.

[4] Ge Z., Liu S., Wang F., Li Z., Sun J., (2021). Yolox: Exceeding yolo series in 2021. *arXiv preprint arXiv*:2107.08430, Japan.

[5] Li R., Huang Y., Wei L., Qin R., Li M., (2021). Design of large-scale intelligent pig system based on RFID (基于 RFID 的规模化智能养猪系统设计). *Modern Agricultural Equipment*. 42(04): 60-63, Guangxi/China.

[6] Lin T Y., Goyal P., Girshick R., He K., Dollár P., (2017). Focal loss for dense object detection. *Proceedings of the IEEE international conference on computer vision,* pp. 2980-2988, United States.

[7] Liu W., Anguelov D., Erhan D., Szegedy C., Reed S., Fu C Y., Berg A C., (2016). SSD: Single shot multibox detector. *European conference on computer vision,* pp. 21-37, Germany.

[8] Maselyne J., Nuffel A V., BD Ketelaere., Bart D K., Jürgen V., Engel F. Hessel C., Bart S., Wouter S., (2014). Range measurements of a high frequency radio frequency identification (HF RFID) system for registering eating patterns of growing-finishing pigs. *Computers & Electronics in Agriculture,* 108 (10): 209-220, Beijing/China.

[9] Qin Q., Liu Z., Zhao C., Zhang C., Dai D., Sun J., Wang Z., Li J., (2021). Application of machine vision technology in animal husbandry (机器视觉技术在畜牧业中的应用). *Agricultural Engineering,* 11(7): 27-33, Neimenggu/China.

[10] Ren S., He K., Girshick R., Sun J., (2017). Faster R-CNN: towards real-time object detection with region proposal networks. *IEEE Transactions on Pattern Analysis & Machine Intelligence,* 39(6): 1137-1149, United States.

[11] Statham P., Hannuna S., Jones S., Campbell N., Mendl M., (2020). Quantifying defence cascade responses as indicators of pig affect and welfare using computer vision methods. *Scientific Reports,* 10(1), United States.

[12] Shen M., Liu L., Yan L., Lu M., Yao W., Yang X., (2014). Review of Monitoring Technology for Animal Individual in Animal Husbandry (畜禽养殖个体信息监测技术研究进展). *Transactions of the Chinese Society for Agricultural Machinery,* 45(10): 245-251, Nanjing/China.

[13] Tan M., Pang R., Le Q V., (2020). Efficientdet: Scalable and efficient object detection. *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition,* pp. 10781-10790, Germany.

[14] Wang L., (2021). Design of Intelligent and Precise Livestock Breeding Management System—Taking Pig Breeding as an example (智能精准畜禽养殖管理系统设计——以猪的养殖为例). *Journal of Shandong Institute of Commerce and Technology,* 21(02): 107-110, Shandong/China.

[15] Wu S., Bao Y., Chen G., Chen Q., (2020). Contactless Identification System for Pig Behavior Based on Machine Vision (基于机器视觉技术的猪行为活动无接触识别系统). *Computer Systems and Applications,* 29(04): 113-117, Guizhou/China.

[16] Yang Q., Xiao D., Zhang G., (2018). Automatic Pig Drinking Behavior Recognition with Machine Vision (猪只饮水行为机器视觉自动识别). Transactions of the Chinese Society for Agricultural Machinery, 49(06): 232-238, Guangdong/China.