ISSN 2616-5775 Vol. 6, Issue 12: 20-32, DOI: 10.25236/AJCIS.2023.061203

Detection of ear root feature region in heterogeneous images based on pig head posture

Qi Li, Shuanglin Nie*

School of Electronic Information and Artificial Intelligence, Shaanxi University of Science & Technology, Xi'an, 710021, China
2655151607@qq.com
**Corresponding author*

Abstract: *In the scene of pig thermal infrared temperature measurement, the temperature of the ear root is closest to the body temperature of the pig. In order to address the problem of inaccurate positioning of the ear root area in the thermal infrared image of the pig under motion due to changes in the pig's head posture and single pixel information, based on the research of automatic inspection and temperature measurement robots, it is proposed to establish an ear root area detection model through visible light images, and map the detection results of the model to the thermal infrared image with the same contour information after registration, achieve accurate localization of the pig ear root area in the thermal infrared image. This algorithm achieves the detection of key points in pig heads in visible light images by improving the YOLOv8 network. By analyzing the position information of key points in pig heads through light shadow projection and dynamic coordinate system adjustment methods, the pig head posture and adjustment angle are obtained. The visible light image of the pig's head after angle adjustment is redivided into ear root regions to correct the impact of pig movement on ear root region division. By using AKAZE and GMA algorithms to extract and register features from heterogeneous images, visible and thermal infrared images with the same contour and scale information are obtained. Finally, the detection results of the ear root region in the visible light image are mapped to the thermal infrared image. Through testing on 2650 datasets, the accuracy of key point and category detection was 99%, the number of key point detection model parameters decreased by 41%, and the registration error was only 1.6%. The accuracy of ear root localization in heterogeneous images reached over 99%, making it easy to achieve accurate temperature measurement.*

Keywords: *Pig ear root area, Key point detection, Pig head posture, Heterogeneous image registration*

1. Introduction

Monitoring pig body temperature in the pig breeding industry is of great significance for assessing the health status, disease warning and control, improving breeding production efficiency, fine management, and quality control of pigs[1-3].

The normal body temperature of adult pigs ranges from 38 to 39.5 ℃. The traditional method of measuring pig body temperature mainly involves manually measuring the rectal temperature of pigs through a thermometer (3-5 minutes), which is inefficient and prone to causing stress reactions in pigs[4]; Some use handheld animal infrared thermometers for temperature measurement, which has low automation and is prone to cross infection between humans and animals. The existing pig temperature measurement methods have certain limitations in terms of contact interference, labor time cost, and real-time performance.

With the advancement of technology, infrared thermal imaging technology has become an advanced non-contact, real-time temperature monitoring method[5]. Previous experimental studies have shown that the ear root temperature of live pigs is closest to the anal surface temperature[6-7], with an error of only 0.26 ℃. Therefore, accurately identifying and locating the characteristic areas of pig ears is the key to using infrared thermal imaging technology for pig body temperature detection.

In the study of obtaining regions of interest in pig thermal infrared images, Zhu Weixing et al. proposed an active shape model based ear region detection method for pigs[8], which can extract feature regions of pig ears to a certain extent. However, when the pig posture changes too much, it is easy to cause detection failure. Ma Li et al. proposed a method based on skeleton scanning strategy for detecting

ISSN 2616-5775 Vol. 6, Issue 12: 20-32, DOI: 10.25236/AJCIS.2023.061203

the surface temperature of pig ear roots using FDSST[9]. When there is significant deviation in the head posture of the pig, it can cause inaccurate ear root positioning, leading to temperature measurement deviation. Zhou Liping et al. proposed an improved Otsu algorithm based ear root feature region detection for pig thermal infrared images[10], which cannot accurately detect the ear root region of pig heads under dynamic conditions. Liu Gang et al. proposed a thermal infrared video detection method for pig ear root temperature based on improved YOLO v4[11-12], which removes skewed frames of pig posture while retaining upright frames, which can easily cause information loss. Zhao Haitao proposed a pig body temperature detection and key temperature measurement part recognition based on infrared thermal imaging technology. The key point detection technology is used to automatically recognize the key parts of the pig face infrared image[13]. This method is suitable for static pig face data and is difficult to combine with automated equipment for real-time detection.

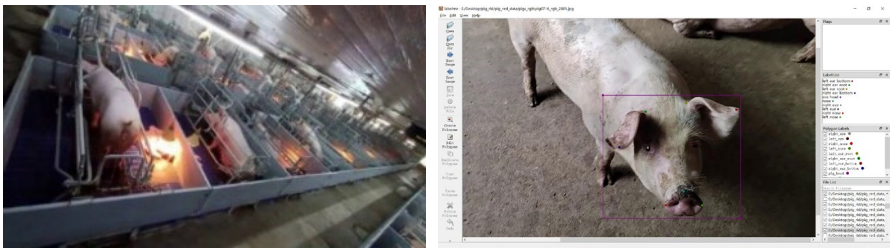
In response to the above issues, this article proposes a pig ear root region detection algorithm based on head posture for heterogeneous images of pigs in Xinda Animal Husbandry, Henan Province, on the basis of a built pig house inspection temperature measurement robot. By improving the YOLOv8 model to detect the head posture of pigs in visible light images, it eliminates obstruction interference and corrects ear root region misdivision under skewed posture; Registering the thermal and visible light images of pigs obtained simultaneously, ensuring that they have the same contour and scale information, mapping the detection results of the ear root area in the visible light image to the thermal image, achieving convenient and accurate positioning of the pig ear root feature area in the thermal infrared image with rich pixel information, which is conducive to accurately obtaining the pig body temperature, To meet the demand for temperature monitoring in modern pig farming.

2. Materials and Methods

2.1. Dataset Creation

The experimental pig data collection location is the fattening house of Henan Xinda Animal Husbandry Co., Ltd. for large white pigs. In order to solve the problem that the image acquired by a single acquisition device may lead to insufficient stability and adaptability of the light detector, the HONOR50 100 million pixel mobile camera and the Jierui Wetong 200W pixel industrial camera are used as the acquisition devices for visible light data, with resolutions of 1920 * 1080 and 640 * 480 respectively. The thermal infrared image is collected using a lepton3.0 infrared camera with a resolution of 160 * 120. A total of 100 visible light and thermal infrared videos were collected from a single pig in the restricted area and a total of 50 live pigs in complex backgrounds. The collection environment is shown in Figure 1a). Finally, a dataset of 2650 pig head images was formed.

Use the Labelme tool to label the target box and key points on the collected pig dataset. The target box is the pig head area (pig_head). In order to determine the pig head posture and ear root area based on the position relationship of the key points, we define the pig head key points to include: left ear tip (left_ear_bottom), right ear tip (right_ear_bottom), left ear root (left_ear_root), right ear root (right_ear_root), left eye (left_eye), right eye (right_eye), left nasal wing (left_nose) Right nasal wing (right_nose), nose center (nose). The annotation of key points on the head of a live pig is shown in Figure 1b).



a)Collection environment. *b)Annotation of pig head key point dataset.*

Figure 1: Dataset collection and annotation.

After completing the annotation of the dataset, in order to improve the robustness and generalization of the model, the original annotated dataset was subjected to data augmentation operations such as rotation, cropping, scaling, brightness adjustment, Gaussian noise addition, and occlusion, resulting in 13250 datasets. The enhanced dataset was divided into three parts in a 7:2:1 ratio: 9275 training sets, 2650 validation sets, and 1325 test sets.

ISSN 2616-5775 Vol. 6, Issue 12: 20-32, DOI: 10.25236/AJCIS.2023.061203

2.2. Construction of pig head key point detection model

Key point detection methods include overall methods, constrained local model methods (CLM), and regression based methods. The overall method has limited generalization ability, making it difficult to fit invisible changes in the pig's head and resulting in low detection accuracy; The constrained local model method has higher robustness to the ambient lighting of pig houses and pig head occlusion, but there is a trade-off between accuracy and robustness; A regression model trained based on regression dependence can quickly and accurately detect key points in different images, with good robustness and generalization ability. This includes direct regression, cascade regression, and deep learning regression. The deep learning regression method is superior to other methods in terms of detection accuracy and speed, and is only time-consuming during the training process. Therefore, we chose to construct a pig head key point detection model based on deep learning regression method.

The YOLO series of deep learning algorithms are superior to similar algorithms. YOLOv8 supports omnidirectional visual AI (Artistic Intelligence) tasks, including detection, segmentation, classification, and pose estimation. YOLOv8 provides a new SOTA (state of the art) model based on YOLOv5. Replace the C3 module with a C2f module with richer gradient information in the backbone network section; The Head section uses the current mainstream decoupling head to separate the classification and detection heads, and uses the Anchor Free method; The loss function uses a combination of classification BCE and regression CIOU+VFL; The box matching strategy has been changed from static matching to dynamic matching (Task Aligned Assignor). However, the original YOLOv8 network required a larger model to ensure the accuracy of key point detection, making it difficult to deploy on small devices. Therefore, this article improves the original YOLOv8 network by introducing a lightweight Slim Neck network structure[14] to reduce model parameters, improve prediction speed, and add an EMA attention mechanism[15] to improve the model's detection accuracy of key points in pig heads through cross spatial learning strategies. The improved YOLOv8 network is shown in Figure 2.

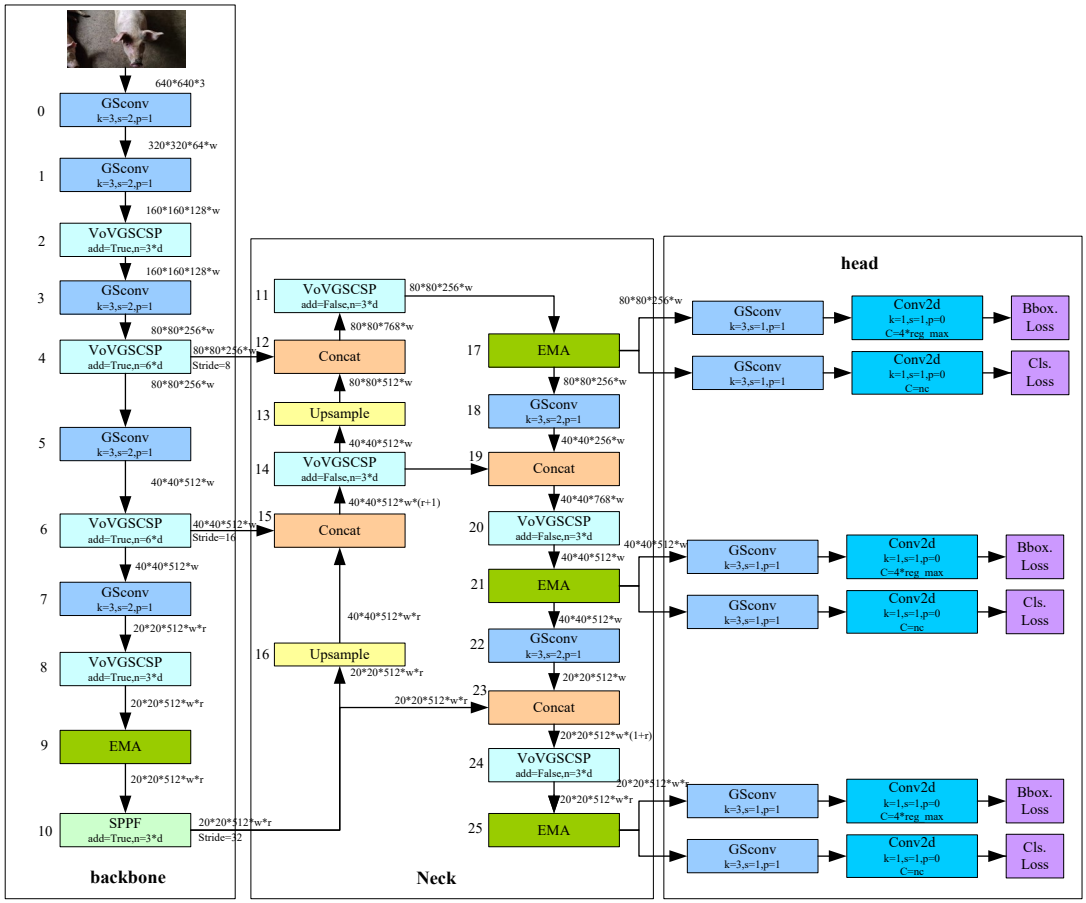


Figure 2: Improved YOLOv8 network structure diagram.

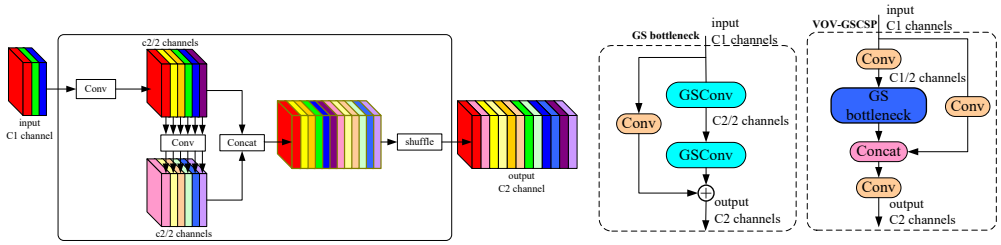
2.2.1. Introduction of lightweight network structure SlimNeck

The SlimNeck network structure includes two lightweight structures: guided spatial context (GSConv)

ISSN 2616-5775 Vol. 6, Issue 12: 20-32, DOI: 10.25236/AJCIS.2023.061203

and one-time aggregated cross level partial network (VoV-GSCSP). Among them, GSConv combines Standard convolution (SC) to capture global information of input feature maps and Depth separable convolution (DSC) to significantly reduce model complexity through point convolution. By reducing the number of model parameters and floating-point operations, GSConv outperforms the original model in accuracy and speed, with a computational cost of approximately 60% to 70% of SC. The structure of GSConv is shown in Figure 3a). The VoV-GSCSP module not only reduces the time complexity of the computational process and network structure, but also improves the model's expression ability and maintains sufficient accuracy. Therefore, using VoV-GSCSP instead of the feature extraction module (C2f) will reduce FLOPs by an average of 15.72% compared to the latter. The VoV-GSCSP module structure diagram is shown in Figure 3b).

Improved YOLOv8 network, introducing SlimNeck network structure. Firstly, GSConv was used to replace all SC operations in the original network head and backbone sections. Secondly, VoV-GSCSP was used to replace all C2f modules, resulting in a 41% reduction in model parameter count and a 1.4ms increase in prediction speed per sheet.

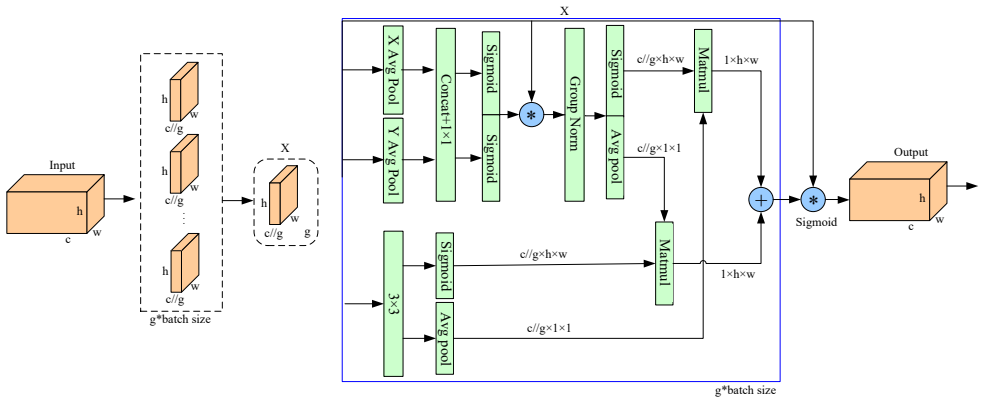


a) GSConv. *b) VoV-GSCSP.*

Figure 3: Slimneck network related modules.

2.2.2. Introducing EMA Attention Mechanism

EMA (Efficient Multi Scale Attention) is an efficient multi-scale attention module based on cross spatial learning. Firstly, by grouping features, some channel dimensions are reshaped into batch dimensions, and channel dimensions are grouped into multiple sub features to evenly distribute spatial semantic features within each feature group, thereby preserving information on each channel and reducing computational overhead. Secondly, a cross spatial learning strategy is used to construct local cross channel interactions for each parallel sub network and fuse the attention maps of two parallel sub networks, fusing contextual information of different scales, enabling deep convolutional neural networks to generate better pixel level attention for higher-order feature maps[16]. The structure of the EMA module is shown in Figure 4. Improve YOLOv8 network by adding EMA attention mechanism in the backbone and head sections to enhance object detection capability.

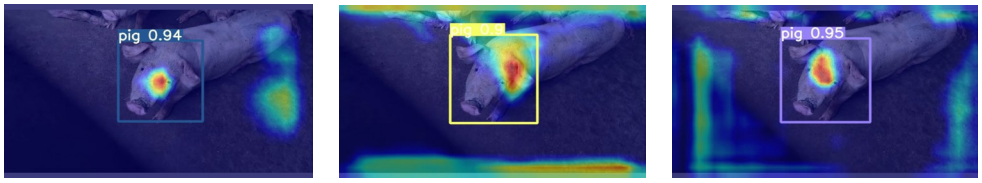


Note: "g" represents the divided group, "X average pool" represents the 1D horizontal global pool, and "Y average pool" represents the 1D vertical global pool, respectively.

Figure 4: EMA attention mechanism module.

In order to better demonstrate the effectiveness of the EMA attention mechanism, the same layer detection results of the original YOLOv8 network and the YOLOv8 network with GCT and EMA attention mechanisms added respectively were compared through thermal graph visualization[17-18], and the results are shown in Figure 5. As shown in the figure, the network with the addition of EMA mechanism significantly improves the accuracy of detection categories and locations.

ISSN 2616-5775 Vol. 6, Issue 12: 20-32, DOI: 10.25236/AJCIS.2023.061203



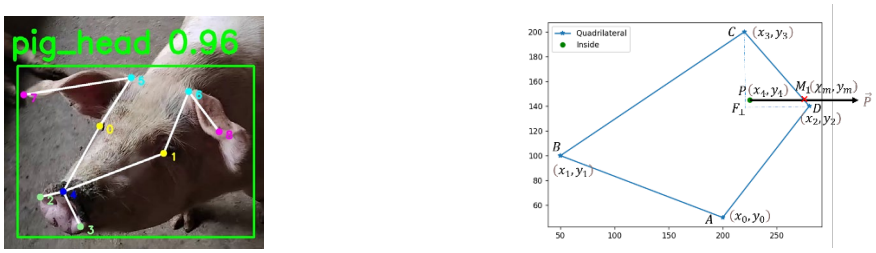
a)YOLOv8 detection result. b)Add GCT attention mechanism detection result. c)Add EMA attention mechanism detection result.

Figure 5: Visualization results of thermal diagram.

2.3. Head pose detection and ear root area adjustment of live pigs

Due to the fact that the camera of the pig house inspection and temperature measurement robot built in the early stage is installed horizontally downwards, it is necessary to focus on studying the effects of raising and lowering the head on locating the ear root area, in order to determine whether the ear root is visible and adjust the position of the ear root area in the future. After completing the task of accurately extracting pig head key points through the pig head key point detection model, it is necessary to analyze the information of pig head key points to obtain the pig head posture. Through analyzing the collected 2650 pig head datasets and conducting research on pig life habits[19-20], it can be seen that when pigs lower their heads, the positions of their eyes and nose are downward, and the angle between the line between the lowest point of their mouth and the neck and the horizontal direction of the neck exceeds -10°, At this point, the ear root area of the pig is clearly visible, which is a 30 * 20 pixel neighborhood centered on the ear root key point (ear_root); When a pig raises its head, its eyes and nose are facing upwards, and the angle between the line connecting the lowest point of its mouth and the neck and the horizontal direction of the neck exceeds+10°. At this point, the ear root area of the pig will be blocked by the face or deviate from the key point position. Based on the above rules, a pig head pose detection method based on light shadow projection algorithm is proposed. When detecting the ear root area of pigs, only the ear root area of pigs with a head down and head up deviation angle less than 10° is detected to eliminate misdivision of the ear root area caused by occlusion interference.

The detection results and numerical identification of key points on the pig head are shown in Figure 6a), where the right eye, left eye, right nasal wing, left nasal wing, and nasal center are marked as 0, 1, 2, 3, and 4, respectively. Define the quadrilateral shape (0, 1, 3, 2) formed by the eyes and bilateral nasal wings as the pose detection polygon ABCD, and the ray starting from the nasal center point is defined as P. Form a right triangle with the four oblique edges of ABCD. If P intersects with the edges of ABCD, the intersection point is denoted as Mi (i=0,1,2), where i represents the i-th intersection point. The extracted pose detection polygon and nose center point are shown in Figure 6b).



a)Key Points and Digital Identification of Pig Head. b)Pose Detection Polygon and Nose Center Point.

Figure 6: Simplified diagram of key point model for pig head.

Using the light shadow projection algorithm to determine the relative position relationship between the nose center point P and the posture detection polygon ABCD, the pig's head posture is determined: if the nose center point is located inside the posture detection polygon, it is determined as a head up; if the nose center point is located on the edge or outside of the posture detection polygon, it is determined as a head down. Taking the CD edge in Figure 6b) as an example, the relevant calculation formula is as follows. From equation (1), equation (2) can be obtained, and equation (3) is the attitude determination result. Determine the number of intersection points between ray P and all edges of the attitude detection polygon ABCD. If it is an even number, point P is outside ABCD; if it is an odd number, point P is on or inside ABCD.

ISSN 2616-5775 Vol. 6, Issue 12: 20-32, DOI: 10.25236/AJCIS.2023.061203

$$\begin{cases} \angle FCD = \theta \\ \tan\theta = \dfrac{|FD|}{|FC|} = \dfrac{|x_D - x_C|}{|y_D - y_C|} \\ \tan\theta = \dfrac{L}{|y_C - y_P|} \\ |x_M - x_C| = L * \tan\theta \\ x_M = x_c + L \end{cases} \tag{1}$$

$$x_M = x_c + \frac{|x_D - x_C|}{|y_D - y_C|} * |y_C - y_P| \tag{2}$$

$$\begin{cases} x_P > x_M, \text{ Point P is located outside polygon ABCD}->Look\ \ down \\ x_P \le x_M, \text{ Point P is located within polygon ABCD}->Look\ \ up \end{cases} \tag{3}$$

Due to the fact that detection algorithms typically use rectangular boxes when identifying the ear root area, if the pig's head posture is not regular, it will result in the divided area including non ear root areas such as forehead and ears, deviating from the accurate ear root area. Calculate the head deviation angle of pigs in visible light images using a dynamic coordinate system, adjust the angle of real-time images with head deviation angle greater than 7 °, and correct the ear root area divided under posture deviation. The calculation steps for live pigs head deviation angle based on dynamic coordinate system are as follows, and the principle of ear root area adjustment is shown in Figure 7.

(1)Establish a dynamic coordinate system with the midpoint O of the pig's head's eyes as the coordinate origin. The horizontal right direction of this point is the positive x-axis direction, and the vertical downward direction is the positive y-axis direction;

(2)Calculate the angle $\angle noy$ between the midpoint of the nose n and the y-axis. If $\angle noy < 7°$, the posture is correct and does not need to be adjusted. Otherwise, it is determined that the head is tilted and the image needs to be rotated according to the size of the angle;

(3)Determine the direction of deviation. If $x_n < x_o$, the head is tilted to the right and needs to rotate counterclockwise. If not, the head is tilted to the left and needs to rotate clockwise.

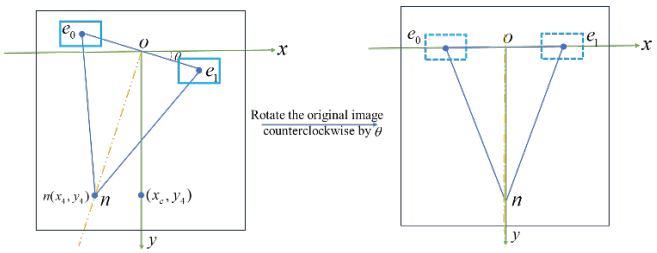


Figure 7: Principle of adjusting the skew head pose.

2.4. Heterologous image registration based on contour and feature point information

Thermal infrared images are essentially grayscale images, and different temperatures cause differences in grayscale values. Visible light images have richer color and texture information compared to thermal infrared images. For visible light images (640 * 480) and thermal infrared images (160 * 120) with different resolutions, the former has a larger viewfinder range than the latter. Registering these two real-time images of live pigs[21-22] allows for the localization of the ear root region through visible light images, and mapping the detection results to the thermal infrared image. This avoids the problem of false or missed detection of the ear region during live pigs movement due to the limited information in the thermal infrared image.

Firstly, gamma transform is used to enhance the contrast of heterogeneous images. Furthermore, Canny edge detection algorithm is used to extract pig contour information from the thermal infrared

ISSN 2616-5775 Vol. 6, Issue 12: 20-32, DOI: 10.25236/AJCIS.2023.061203

image as a template image. Based on the template image, scale transformation is performed on the visible light image of the pig, thereby unifying the size of spatial objects on the heterogeneous image and obtaining the optimal matching position of the thermal infrared image in the visible light image, Then, the AKAZE (Accelerated KAZE) feature extraction algorithm is used to find the top 10% feature points among the 600 feature points in the heterogeneous image[23-25]. GMS (Grid based Motion Statistics) is used to optimize the matching pairs and perform image alignment to complete registration[26-29]. The registration process is shown in Figure 8. The specific implementation of AKAZE and GMS registration is as follows.

(1)The AKAZE algorithm uses nonlinear diffusion filtering to denoise the noise generated by pig head movement and environmental light changes, while preserving edge information, better preserving the edges and details of pig images. The diffusion coefficient varies continuously based on the image gradient obtained during each iteration. The partial differential equation for nonlinear diffusion filtering is shown in equation (4), where div and ∇ represent divergence and gradient, respectively. The diffusion time t is the scale parameter, and the larger t, the lower the complexity of the image content. $c(x,y,t)$ is a conduction function that determines whether the degree of diffusion is suitable for the local structure of the image. In equation (5), L_σ is a Gaussian smooth image, ∇L_σ is the gradient of image L_σ. The function g has three structures, such as equations (6), (7), and (8), where k is the alignment factor that controls diffusion. g1 prioritizes the preservation of high contrast edge parts, g2 preserves areas with larger image widths, and g3 smooths the internal areas of the image while preserving edge information. This article uses the g2 function to better obtain edge information in pig images. Equation (9) represents the corresponding step size.

$$\frac{\partial L}{\partial \mathrm{t}} = div(c(x,y,t)\bullet\nabla L) \tag{4}$$

$$c(x,y,t) = \mathrm{g}(\left|\nabla L_\sigma(x,y,t)\right|) \tag{5}$$

$$\mathrm{g}_1 = \exp(-\frac{\left|\nabla L_\sigma\right|^2}{k^2}) \tag{6}$$

$$g_2 = \frac{1}{1+\frac{\left|\nabla L_\sigma\right|^2}{\mathrm{k}^2}} \tag{7}$$

$$g_3 = \begin{cases} 1, & \left|\nabla L_\sigma\right|^2 = 0 \\ 1-\exp\left(-\frac{3.315}{(\left|\nabla L_\sigma\right|/k)^3}\right), & \left|\nabla L_\sigma\right|^2 > 0 \end{cases} \tag{8}$$

$$\tau_j = \frac{\tau_{\max}}{2\cos^2(\pi\frac{2j+1}{4n+1})} \tag{9}$$

The image obtained through nonlinear filtering by solving equations (4), (5), (7), and (9) is:

$$L^{(i+1,j+1)} = (I+\tau_j A(L^i))L^{(i+1,j)} \mid j = 0,1,...,n-1 \tag{10}$$

(2)AKAZE achieves feature point detection in heterogeneous images by searching for normalized Hessian local maximum values at different scales. The Hessian matrix calculation formula is:

$$L^i_{Hessian} = \sigma^2_{i,norm}(L^i_{xx}L^i_{yy} - L^i_{xy}L^i_{xy}) \tag{11}$$

In the formula, $\sigma_{i,norm}$ is the normalized scaling factor; L^i_{xx} and L^i_{yy} are second-order transverse and longitudinal differentials, respectively; L^i_{xy} is a second-order cross differentiation.

(3)AKAZE uses M-LDB descriptors for the detected live pigs image feature points. This descriptor

ISSN 2616-5775 Vol. 6, Issue 12: 20-32, DOI: 10.25236/AJCIS.2023.061203

divides the neighborhood into n*n square grids, samples the pixels within the grid, and calculates the mean grayscale of all sampled pixels in the grid, as well as the first derivative of each pixel in the x and y directions. The sampling interval is determined by the scale parameters σ of the feature points, which not only reduces the computational complexity but also makes the descriptor more robust to scale changes. The relevant calculation formula is as follows. Among them, $I_{\text{avg}}(i)$ is the average strength of each grid unit, $d_x(i)$ is the gradient in the x direction, and $d_y(i)$ is the gradient in the y direction. Perform binary testing on the grid using equation (15) and encode to obtain a binary descriptor (16).

$$I_{\text{avg}}(i) = \frac{1}{m}\sum_{k-1-m} Intensity(k) \tag{12}$$

$$d_x(i) = Gradient_x(i) \tag{13}$$

$$d_y(i) = Gradient_y(i) \tag{14}$$

$$F(i) = \{I_{\text{avg}}(i), d_x(i), d_y(i)\} \tag{15}$$

$$\tau(F(i), F(j)) = \begin{cases} 1, & F(i) - F(j) > 0, i \neq j \\ 0, & other \end{cases} \tag{16}$$

(4)GMS divides the image into multiple grids and performs feature matching within each grid, thereby utilizing the spatial information within the grid to improve the quality of feature matching. Measure the quality of feature matching using the statistical likelihood of each grid matching frequency, and obtain the optimal feature matching by maximizing the statistical likelihood of all grids. For the matching point pairs{x, y} in the grid, their positive and negative values can be determined using equation (19). Equation (19) can be obtained from equations (17) and (18). In equation (17), K is the number of small domains near the matching point, $\chi_{a^k b^k}$ is the number of matching pairs in the two matching domains, and S_i basically follows a binomial distribution. In equation (18), n is the average number of feature points in each small domain, and P is the evaluation function for S_i right and wrong judgment ability. s_t and s_f are the standard deviations of matched positive and erroneous S_i, while m_t and m_f are the mean values of matched positive and erroneous S_i, respectively.

$$S_i = \sum_{k=1}^{K} \left| \chi_{a^k b^k} \right| - 1 \tag{17}$$

$$P = \frac{m_t - m_f}{s_t - s_f} = \sqrt{Kn}\frac{p_t - p_f}{\sqrt{p_t(1-p_t)} + \sqrt{p_f(1-p_f)}} \tag{18}$$

$$\{i, j\} \in \begin{cases} T, & Sij > \tau_i = a\sqrt{n_i} \\ F, & other \end{cases} \tag{19}$$

(5)Use optimized matching to align the obtained homography matrix with heterogeneous images.

ISSN 2616-5775 Vol. 6, Issue 12: 20-32, DOI: 10.25236/AJCIS.2023.061203

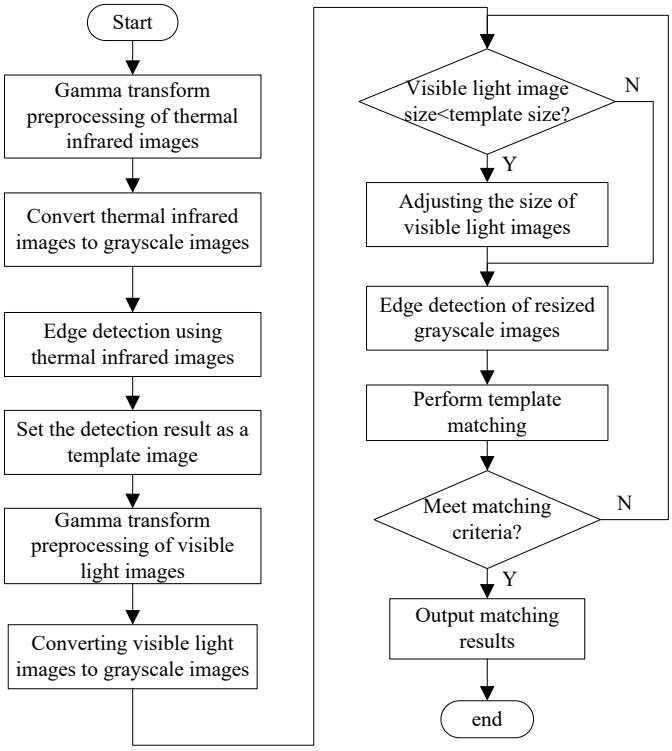


Figure 8: Registration flow.

3. Results and Analysis

3.1. Evaluation of the effectiveness of the pig head key point detection model

In order to objectively evaluate the performance of the improved pig head key point detection model[30], the original model and two improved models were tested using a self built live pigs dataset on a server configured with NVIDIA GeForce RTX 2080 Ti graphics card. Precision (P), recall (R), average precision (AP_{50}) The average accuracy (AP_{IOU}) and calculation parameters (parameters) when the IOU threshold gradually increases between 0.5 and 0.95 are used as evaluation indicators to evaluate the prediction box and key prediction points. The calculation formula for the indicators is as follows.

$$P = \frac{T_P}{T_p + F_P} \tag{20}$$

$$R = \frac{T_P}{T_p + F_N} \tag{21}$$

$$AP_{50} = \int_0^1 P(R)dR \tag{22}$$

$$AP_{IOU} = \frac{1}{N}\sum_{i=1}^{N} A_{p_i} \tag{23}$$

In equations (20) and (21), TP represents the correctly predicted number of live pigs' heads, FN represents the number of missed pig heads, and FP represents the number of false pig heads. The area below the PR curve composed of accuracy and recall is defined as AP_{50} in equation (22), where N represents the total number of categories and APi represents the detection accuracy of class i. The test results are shown in Table 1. From the table, it can be seen that the improved SlimNeck+EMA model in this article not only achieves a compression rate of 41% for calculation parameters and an inference speed increase of 1.4ms per image, but also improves all evaluation indicators. The P, R, and AP_{IOU} of the prediction box have increased by 4.95%, 0.1%, and 0.3%, respectively. The P, R, AP_{50} and AP_{IOU} of key

ISSN 2616-5775 Vol. 6, Issue 12: 20-32, DOI: 10.25236/AJCIS.2023.061203

points have increased by 4.95%, 0.1%, 0.2%, and 0.1%, respectively.

Table 1: Test results of three models.

Network parameters	Box(P	R	AP_{50-95})	Pose(P	R	AP_{50}	AP_{50-95})	P(MB)	FPS
YOLOv8	0.949	0.999	0.983	0.949	0.999	0.993	0.993	139.4	117
SlimNect+GCT	0.991↑	0.973↓	0.986↑	0.991↑	0.973↓	0.994↑	0.993↑	81.1↓ (-41%)	121
SlimNect+EMA	0.996↑	1↑	0.986↑	0.996↑	1↑	0.995↑	0.994↑	81.5↓ (-41%)	140

The detection effects of three types of networks on the same sample are shown in Figure 9. As shown in the figure, the network with the addition of EMA attention mechanism has more accurate regression box and key point detection results for multiple live pig targets. Suitable as a key point detection model for live pigs head posture analysis.

a)YOLOv8 detection effect. b)SlimNect+GCT detection effect. c)SlimNeck+EMA detection effect.

Figure 9: Three network detection results.

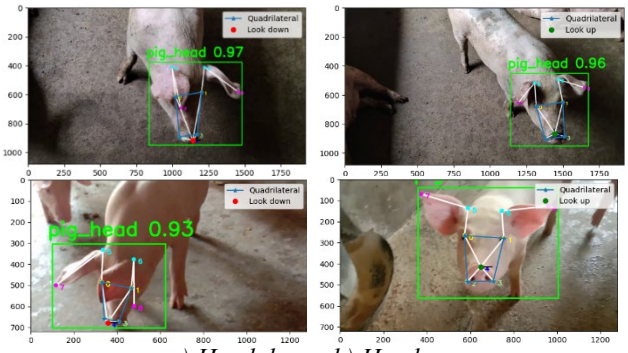
3.2. Evaluation of pig head posture detection and ear root area adjustment effect

The detection results of four pig head postures and adjustment numbers on 1000 visible light datasets are shown in Table 2. Among them, the accuracy of judging the head up and head down posture of live pigs reached 100%, and adjustments were made to the side face posture with excessive deviation angle. The detection results show that the light shadow projection algorithm has a high accuracy in detecting the head posture of live pigs, and the adjustment rate for posture with excessive deviation in the side face posture reaches 80%, which is conducive to correcting the ear root region division.

Table 2: Calculation results of pig head posture detection and adjustment angle.

test sample	Total number	Head up	Head down	Adjust number	Adjust scale	Recognition accuracy
Head up & Lateral face	125	125	0	100	80%	100%
Head up & face up	86	86	0	0	0	100%
head down &Lateral face	473	0	473	380	80%	100%
Head down & face up	316	0	316	0	0	100%

The head posture detection results of some live pigs are shown in Figure 10.

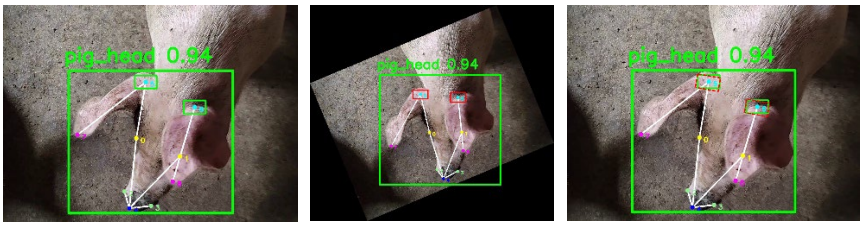


a) Head down. b) Head up.

Figure 10: Detection results of pig head posture.

ISSN 2616-5775 Vol. 6, Issue 12: 20-32, DOI: 10.25236/AJCIS.2023.061203

The results of adjusting the head posture and dividing the ear root region of some live pigs are shown in Figure 11. It can be seen from Figure 11 c) that the adjusted division of the ear root area of live pigs reduces the interference of their ears and forehead areas, making the ear area division more reasonable and accurate.



a) Original head deviation detection results. b) Adjusted detection results. c) Comparison of ear root area division.

Figure 11: Effect of ear root area detection after adjustment.

3.3. Evaluation of ear root feature area detection effect

Under the changing conditions of light and breeding environment, 1000 images of single and multiple pigs with different postures were registered using visible light and thermal images. The registration results are shown in Table 3, with an average registration error of only 1.6%.

Table 3: Registration error analysis table.

	thermal infrared image	Visible light image	Registration successful	Registration failed	Registration error
Single pig	500	500	494	6	1.2%
Multiple pigs	500	500	490	10	2%

The partial registration process and alignment fusion diagram are shown in Figure 12. The visible light matching image obtained after registration has the same scale and contour information as the thermal infrared image. The accuracy of ear root region localization in the thermal infrared image using visible light images is higher than that of the key point detection model. That is, the accuracy of the method proposed in this article in detecting ear root regions in the thermal infrared image is over 99%.

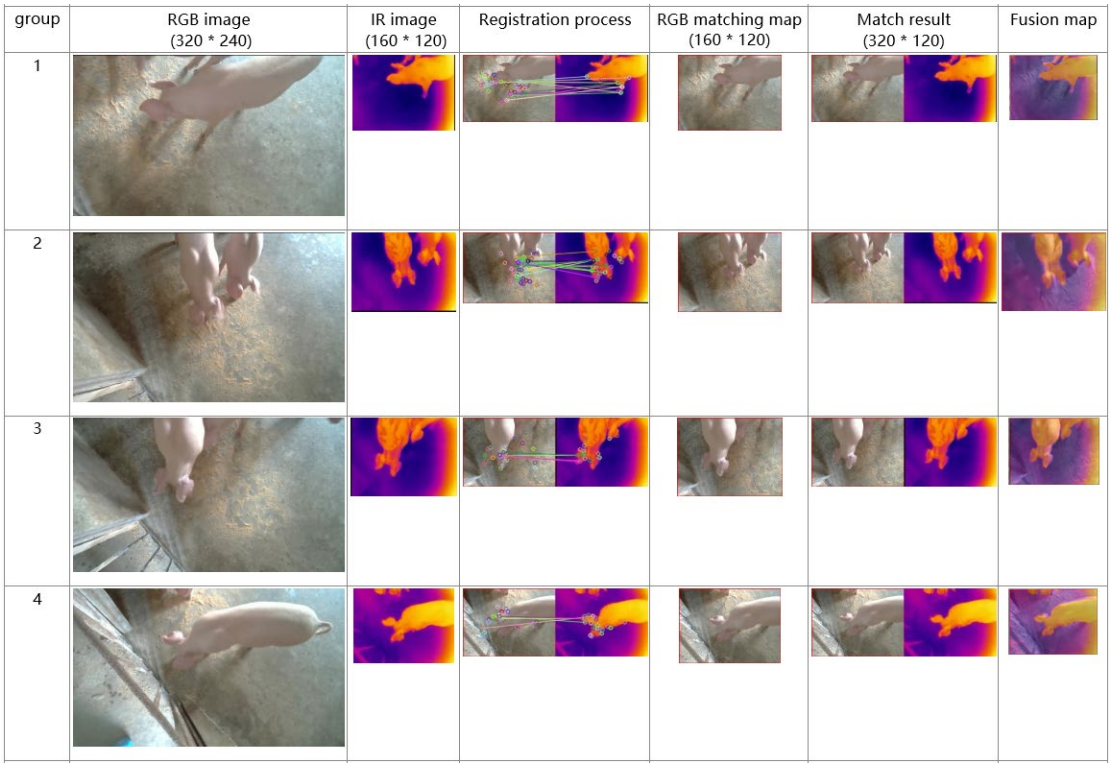
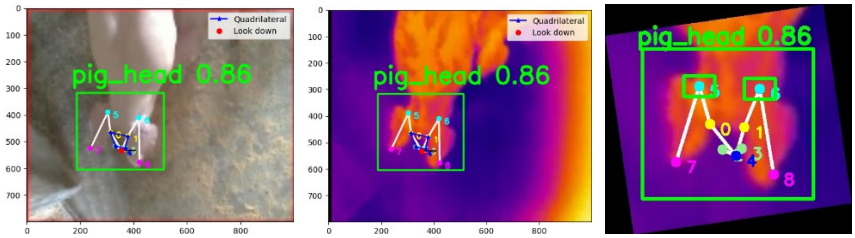


Figure 12: Registration results of visible light image and thermal infrared.

ISSN 2616-5775 Vol. 6, Issue 12: 20-32, DOI: 10.25236/AJCIS.2023.061203

After registration is completed, the ear root region detection results of the visible light registration image are mapped to the thermal infrared image to obtain the ear root key points in the thermal infrared image. After adjusting the head region through the pose adjustment algorithm, the 30 * 20 pixel area around the ear root key points is divided into ear root regions. The final detection results are shown in Figure 13.



a)Original image detection result b)Detection result after mapping c)Ear root area division result after adjusting posture

Figure 13: Ear root region detection results.

4. Conclusions

(1)An improved YOLOv8 pig head key point detection model was proposed, which improved the accuracy of category box and key point detection by introducing EMA attention mechanism. The key point detection accuracy reached 99.4%, and the model was lightweight by introducing SlimNect structure, reducing the number of parameters by 41%, which is conducive to deployment on intelligent AI devices.

(2)A pig head pose detection and ear root area adjustment method based on light shadow projection algorithm and dynamic coordinate system adjustment method was proposed to eliminate ear root misidentification and correction caused by ear root occlusion during pig movement, which is beneficial for accurately locating the ear root area of pigs.

(3)By using a heterogenous image registration algorithm based on contours and feature points, visible light images are registered with thermal infrared images, with an average registration error of 1.6%. The detection results of the ear root area in the visible light image are mapped to the thermal infrared image, achieving accurate positioning of the pig ear root area in the thermal infrared image. Reduce the problem of false or missed detection of the ear root area caused by direct detection of live pig heads under dynamic conditions due to a single thermal image information, and achieve positioning accuracy of over 99%.

References

[1] SOERENSEN D D, CLAUSEN S,MERCER J B, et al. Determining the emissivity of pig skin for accurate infrared thermography[J]. Computers and Electronics in Agriculture, 2014,109:52-58.
[2] SYKES D J, COUVILLION J S, CROMIAK A, et al. The use of digital infrared thermal imaging to detect estrus in gilts[J].Theriogenology, 2012,78(1):147-152.
[3] TZANIDAKIS C,SIMITZIS P,ARVANITIS K, et al. An overview of the current trends in precision pig farming technologies[J]. Livestock Science,2021,249:104530.
[4] KAMMERSGAARD T S, MALMKVIST J, PEDERSEN L J. Infrared thermography-a non-invasive tool to evaluate thermal status of neonatal pigs based on surface temperature[J]. Animal, 2013, 7(12):2026-2034.
[5] ZHANG X, KANG X, FENG N, et al. Automatic recognition of dairy cow mastitis from thermal images by a deep learning detector[J]. Computers and Electronics in Agriculture, 2020,178:105754.
[6] FENG Y Z, ZHAO H T, JIA G F, et al. Establishment of validated models for non-invasive prediction of rectal temperature of sows using infrared thermography and chemometrics[J]. International Journal of Biometeorology, 2019, 63(10):1405-1415.
[7] STUKELJ M, HAJDINJAK M, PUSNIK I. Stress-free measurement of body temperature of pigs by using thermal imaging-useful fact or wishful thinking[J]. Computers and Electronics in Agriculture, 2022, 193:106656.
[8] Zhu Weixing, Liu Bo, Yang Jianjun, et al. Detection method of pig ear region based on improved active shape model [J]. Journal of Agricultural Machinery,2015,46(03):288-295.
[9] Ma Li, Zhang Xudong, Xing Zizheng, et al. FDSST detection method for surface temperature of pig

ISSN 2616-5775 Vol. 6, Issue 12: 20-32, DOI: 10.25236/AJCIS.2023.061203

ear roots based on skeleton scanning strategy [J]. Journal of Agricultural Machinery,2020,51(S1):371-377.
[10] Zhou Liping, Chen Zhi, Chen Da, et al. Ear root feature region detection in pig thermal infrared images based on improved Otsu algorithm [J]. Journal of Agricultural Machinery,2016,47(04):228-232+14.
[11] Liu Gang, Feng Yankun, Kang Xi. A thermal infrared video detection method for pig ear root temperature based on improved YOLO v4 [J]. Journal of Agricultural Machinery,2023,54(02):240-248.
[12] Feng Yankun, Kangxi, Wang Yanchao, et al. Method for detecting the temperature of pig ear roots based on thermal infrared video [J]. Journal of Agricultural Machinery,2021,52(S1):284-290.
[13] Zhao Haitao. Pig body temperature detection and key temperature measurement location recognition based on infrared thermal imaging technology [D]. Huazhong Agricultural University, 2019.
[14] Hulin Li,Hanbing Wei, Zheng Liu et al. Slim-neck by GSConv: A better design paradigm of detector architectures for autonomous vehicles[J]. Computer Vision and Pattern Recognition(CVPR),2022.
[15] D. Ouyang et al., "Efficient Multi-Scale Attention Module with Cross-Spatial Learning," ICASSP 2023-2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Rhodes Island, Greece, 2023, pp. 1-5.
[16] Huang Zhijie, Xu Aijun, Zhou Suyin, et al. A Pig Face Key Point Detection Method Integrating Hyperparameterization and Attention Mechanism [J/OL]. Journal of Agricultural Engineering:1-9[2023-09-01].
[17] Xie Qiuju, Wu Mengru, Bao Jun, et al. Individual Pig Face Recognition by Integrating Attention Mechanism [J]. Journal of Agricultural Engineering,2022,38(07):180-188.
[18] Dongsheng Ruan, Daiyin Wang, Yuan Zheng, Nenggan Zheng, Min Zheng. Gaussian Context Transformer[J]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2021.
[19] Yan Hongwen, Liu Zhenyu, Cui Qingliang, et al. Facial pose detection of group health preserving pigs based on improved Tiny YOLO model [J]. Journal of Agricultural Engineering,2019,35(18):169-179.
[20] Wang Lu, Liu Qing, Cao Yue, et al. Pig herd posture recognition based on improved Cascade Mask R-CNN and collaborative attention mechanism [J]. Journal of Agricultural Engineering, 2023,39 (04): 144-153.
[21] Zhao Yifei. Research on Registration and Fusion Algorithms for Visible and Infrared Images [D]. Xi'an University of Electronic Science and Technology,2021.
[22] Duan Linfeng, Hou Xinguo, Hu Zhiyuan. NSCT contour and principal direction consistency infrared and visible image registration [J]. Electrooptic and control,2022,29(06):1-5.
[23] Zong Huilin, Yuan Xiping, Gan Shu, et al. Improved AKAZE algorithm for feature matching of unmanned aerial vehicle images in debris flow areas [J]. Surveying and Mapping Bulletin,2023(02):91-96+103.
[24] Xue Peng. Research on Substation Augmented Reality Inspection System Based on AKAZE Algorithm [D]. Northeast Petroleum University,2022.
[25] Zhou Rui. Research on Mobile Image Mosaic Algorithm Based on Improved AKAZE-GMS and Grid Optimization [D]. Central China Normal University,2021.
[26] Bian J W, Lin W Y, Liu, Matsushita Y, et al. GMS: Grid-Based Motion Statistics for Fast, Ultra-robust Feature Correspondence. International Journal of Computer Vision, 2020, 128(6):1850-1593.
[27] Hu Xin, Hu Luming, Liu Guihang. An improved GMS image feature point matching algorithm [J]. Electronic measurement technology,2021,44(17):131-137.
[28] Liu Shuai, Rui Ting, Wang Dong, et al. Research on Improved GMS Algorithm Based on Feature Matching Quality [J]. Electro Optics and Control,2021,28(07):31-34.
[29] Xiang Hengyong, Zhou Li, Ba Xiaohui, et al. Feature matching and filtering algorithm based on dynamic window motion statistical information [J]. Journal of South China University of Technology (Natural Science Edition),2020,48(06):114-122.
[30] Li Guangbo, Cha Wenwen, Chen Chengpeng, et al. Pig face recognition and detection method based on improved YOLOv5s [J]. Southwest Agricultural Journal,2023,36(06):1346-1356.