An Improved YOLOv8 Method for Measuring the Body Size of Xinjiang Bactrian Camels

Yue Peng, Alifu Kurban, Mengmei Sang
Department School of Software, Xinjiang University, Urumqi, China

***Abstract*—Camel body size measurement has initially been applied in livestock production. However, current methods suffer from low measurement accuracy due to detection box localization loss and occlusions. This study proposes an effective algorithm, Camel-YOLOv8, specifically designed for detecting Xinjiang Bactrian camels and calculating their body sizes. By integrating the Selective Kernel Networks (SKAttention) mechanism and an enhanced Asymptotic Feature Pyramid Network structure (AFPN-beta), the algorithm successfully captures the body characteristics of Bactrian camels in natural environments and converts these into precise size data. We have developed a Xinjiang Bactrian camel body size measurement dataset and applied the enhanced YOLOv8 model for accurate classification and detection. By extracting key point pixel values and applying Zhang Zhengyou's calibration method, the coordinate system data is converted into accurate body size measurements. The Camel-YOLOv8 achieves a detection accuracy of 76.4% on the Xinjiang Bactrian camel dataset, marking a 3.7% improvement over the baseline model. In terms of body size calculation, the average relative errors for height and chest circumference are -3.39% and 4.1%, respectively, demonstrating high measurement precision. The algorithm not only maintains high detection accuracy but also achieves a reasonable balance between detection speed and efficiency, providing an effective solution for rapid acquisition of camel body size information.**

Keywords*—*YOLOv8; Asymptotic Feature Pyramid Network; SKAttention; Bactrian camel body size measurement

I. Introduction

The scaling and precision enhancement in modern camel farming have led breeders to recognize the critical importance of selecting, cultivating, and fostering the healthy development of superior camel breeds. The physical development and size of camels are reflected through body measurement indicators, closely linked to the camels' adaptability to their environment, productivity, reproductive performance, and economic value [1]. Accurate determination of camel body size and morphological assessments are vital for monitoring camel growth, genetic selection within herds, improving reproductive capabilities, enhancing productivity, and standardizing rearing practices [2]. Traditional livestock measurement and condition assessment methods, which typically involve contact tools like tape measures, can induce stress in animals, impacting their growth and development. With the evolution of technologies such as image processing [3], pattern recognition [4], and artificial intelligence [5], machine learning [6] has increasingly been applied in China's livestock industry, yielding significant results. Leveraging machine vision for measuring camel body size and performing morphological evaluations facilitates a comprehensive analysis of camel development, quickly assessing their growth conditions and preparing for potential health crises.

The application of computer vision [7] in livestock management is diverse and widespread, enhancing management efficiency, productivity, and animal welfare through advanced image analysis techniques. Computer vision systems can autonomously monitor animal activities, behavioral patterns, and body language, identifying behaviors such as eating, resting, and socializing. These insights are crucial for assessing animal health, welfare, and productive performance. By analyzing appearances and behaviors, computer vision technologies can aid in the early identification of disease signs or health issues, such as limping, weight changes, or decreased appetite, enabling timely intervention [8]. Computer vision can also measure animal body dimensions, such as length and height, to estimate weight and monitor growth, aiding in nutritional management and breeding program improvements. By integrating computer vision with automation technologies, feeding quantities and nutritional ratios can be automatically adjusted based on an animal's weight, health condition, and growth needs, supporting individual animal management, population structure analysis, and breeding and selection decisions [9].

The advancement and implementation of computer vision technology have significantly raised the productivity and management levels in the livestock industry, also providing strong support for enhancing animal welfare and ensuring food safety. Through target detection technologies, camel body size data can be calculated using detection boxes. However, ensuring the accuracy of these calculations is crucial, as the precise positioning of detection boxes is paramount. Detection box localization loss refers to inaccuracies in positioning detection boxes during the object detection process due to various factors, impacting the accuracy of body size calculations. Additionally, occlusion among camels, a common issue due to their large size and tendencies to move closely within groups, can cause some parts to be obscured by other camels or objects, affecting the accuracy of detection box positioning. This occlusion can lead to inaccuracies in the detection box area, thereby affecting the results of body size calculations. The primary contributions of this paper include:

1) Establishing a rich dataset containing images of camel postures in various environments, providing necessary data support for algorithm training and evaluation.

2) Current camel body size detection methods suffer from detection box localization loss; in YOLOv8 [10], the

SKAttention mechanism is added to reduce this loss, thereby minimizing errors in calculating camel body size.

3) Using the Asymptotic Feature Pyramid Network (AFPN-beta) structure effectively resolves issues of camel occlusion, ensuring that the body size information of obscured camels can be more accurately calculated.

4) Employing the camera calibration method from Zhang Zhengyou's calibration technique to obtain camera model parameters, such as internal, external, and distortion parameters. Based on the single-camera imaging principle and the transformation relationships between coordinate systems, a measurement model for the body size of Xinjiang Bactrian camels is established, facilitating the measurement of body size information for Bactrian camels.

In summary, Section II of this paper reviews the existing research in the field of camel body measurement and analyzes the current state of computer vision and object detection technologies in livestock management. Section III provides a detailed description of the improved YOLOv8 method proposed in this paper, including feature selection, integration of the SKAttention mechanism, application of the AFPN-beta structure, and the computation methods for camel body metrics. In Section IV, we present the experimental design, dataset preparation, training parameter settings, and evaluation metrics of the model, and provide an in-depth analysis of the experimental results. Section V summarizes the main research findings of this paper and discusses future research directions.

II. Related Work

The use of computer technology for animal biometric measurement through image processing and feature extraction facilitates efficient and accurate measurements, widely applied in biological research and ethology. Federico Pallottino [11] and colleagues compared manual measurements with stereovision measurements, finding a high overall correlation and lower variability, with an average error below 3%. The variability in error magnitude depends on the specific traits, but Pallottino's study did not delve into the theoretical basis of the measurement algorithm. Additionally, the research focused mainly on correlation analysis without involving the design and optimization of specific algorithms, limiting its potential for technical deepening and application expansion. Qin [12] and others introduced deep learning, proposing a unique method for measuring livestock body dimensions using Mask R-CNN [13] to measure cattle and goats of various sizes against different backgrounds. This study initially utilized the idea of human joint location to accurately position livestock body feature points and perform precise measurements. However, due to the unique physiological structure of camels, especially the presence of humps, the use of deep learning models like Mask R-CNN and the search for measurement points can cause significant interference. This issue limits the suitability and accuracy of this technology in camel measurement. Mahdi Khojastehkey [14] and colleagues used SPSS software and Pearson correlation coefficients to select features more relevant to single-humped camel measurement, discovering the mathematical relationships between digital image-extracted features and camel body dimensions. This finding provides an important theoretical basis for developing more accurate camel body dimension calculation models, further supporting the feasibility and effectiveness of using digital image technology for camel body measurement.

Computer vision, a critical component of the field of artificial intelligence, plays a vital role across various industries. With the development of machine learning, especially deep learning, the application of object detection [15], [16] has expanded beyond emerging industries. In traditional sectors, such as livestock, object detection algorithms are gradually showcasing their significant functions. AlNujaidi [17] and others researched a camel-vehicle collision mitigation system through computer vision. Although progress was made in camel detection, the study did not extend to specific calculations of camel body dimensions, limiting its completeness and practicality in biometric applications. Wang [18] and others proposed a portable and automated Xtion measurement system for assessing pig body sizes, which showed significant advantages in measuring pig body dimensions and confirmed the importance of these dimensions in predicting weight. However, this system needs more detailed measurement methods for body dimensions to ensure accuracy when adapting to the more complex morphology of camels. Finally, Li and Teng [19] studied deep learning-based body dimension measurement methods for goats and cattle against different backgrounds. While this method performs well in measuring cattle and goat body sizes, the complexity and specificity of camel body dimensions pose significant challenges for this technology in camel measurement. Wang Yusha [20] and others developed a computer vision-based device for measuring the body dimensions and body mass traits of large yellow croaker, performing well in measuring organisms with smooth body surfaces. However, the same measurement point positioning strategy may not be suitable for camels, which have more structurally complex body surfaces. Munir Ahmad et al. [21] developed a deep transfer learning-based animal face identification model, demonstrating the significant potential of deep learning in the field of biometric identification, particularly for recognizing complex body structures. However, the study primarily focuses on facial recognition and does not fully explore the application of this technology to the measurement of more complex body parts, such as the humps of camels. Additionally, the model's scalability to handle the complex morphological structures of different animal species still requires further validation. Dhivya Mohanavel and Muthu Ishwarya [22] developed a deep learning and computer vision-based animal detection warning system for agricultural environments, highlighting the growing importance of these technologies in agriculture. However, the limitation of this study lies in its primary focus on animal detection and warning functions, without addressing the collection and analysis of specific biometric data. While it provides valuable insights into the application of deep learning in animal management within agricultural settings, the potential for precise animal body measurement and biometric analysis has not been fully explored.

III. Research Method

The Camel-YOLOv8 model primarily consists of three components: the Backbone, the Neck, and the Head networks.

Initially, the input image is processed through multiple convolutional layers (Conv) for preliminary feature extraction. This is followed by several c2f modules that refine the features further. During this process, the SK Attention module enhances the attention mechanism in feature extraction, enabling the model to focus more on important feature regions. Subsequently, features undergo multiscale processing via the SPPF (Spatial Pyramid Pooling - Fast) module. The Neck network then integrates multiple layers of features using convolutional layers and the AFPN-beta (Asymptotic Feature Pyramid Network - beta) module, with further processing by c2f modules. In the Head network, a series of convolutional layers transform the feature map into the format required by the output layer, calculating both the Bounding Box Loss (BBox Loss) and the Classification Loss (Cls Loss). Ultimately, the model outputs the detection results in the image, including bounding boxes and class labels. Overall, the YOLOv8 model, through its multi-layer feature extraction and integration mechanisms, effectively enhances the accuracy and efficiency of object detection while maintaining computational efficiency. The proposed model structure is illustrated in Fig. 1.

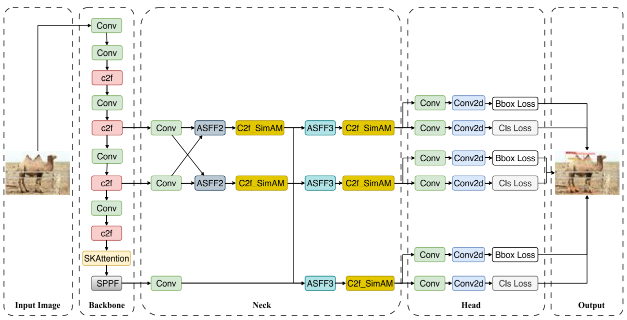

Fig. 1. Structure diagram of camel-YOLOv8 model.

A. Feature Selection

Many research findings indicate that chest girth, body length, pelvic width, and shoulder height are the most suitable and reliable parameters for estimating the live weight of animals. Fig. 2 shows diagram of camel body measurement points.

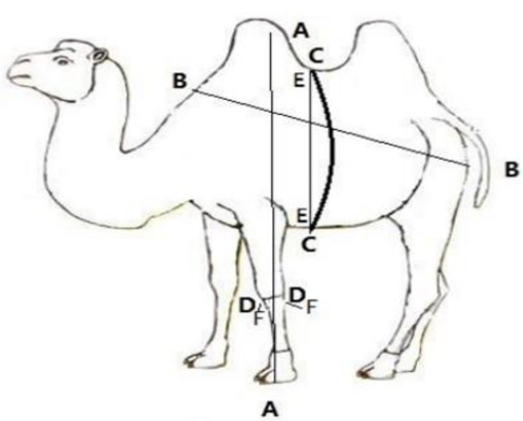

Fig. 2. Diagram of camel body measurement points.

In recent studies, in some cases, features from digital images have been used to estimate the body size of livestock [23]. The camel body measurement points are shown in the diagram. For the Xinjiang Bactrian camel, the height (A) refers to the vertical distance from the base of the camel's front hump to the ground, the body length (B) refers to the distance from the shoulder end to the hip end, the chest girth (C) is the vertical circumference measured from the base of the front hump down through the center of the horny pad at the chest bottom around the body, the cannon circumference (D) refers to the horizontal circumference measured around the cannon at the upper third of the left forelimb, chest diameter (E), and cannon diameter (F).

Using SPSS software, this paper analyzes the correlation between camel weight and other body metrics from existing camel data in Fuyun Town, Fukang City, Xinjiang Uighur Autonomous Region. We have chosen to use the Pearson correlation coefficient [24] to measure the linear relationship between them. The Pearson correlation coefficient is a statistic used to measure the degree of linear correlation between two variables. It is employed to assess the linear relationship between two variables, with values ranging from -1 to +1. A correlation coefficient of 1 indicates a perfect positive correlation; a coefficient of -1 indicates a perfect negative correlation; and a coefficient of 0 indicates no linear correlation between the variables. The formula for calculating the Pearson correlation coefficient is as follows:

$$r = \frac{\mathrm{cov}(X,Y)}{\sigma_X \sigma_Y} \tag{1}$$

Here, Cov represents covariance, and σ represents standard deviation. Specifically, the Pearson correlation coefficient is calculated as the covariance of the two variables divided by the product of their respective standard deviations. Below are the correlation coefficients between camel characteristic information and body weight in the camel data from Ziniquanzi Town, Fukang City, Xinjiang Uighur Autonomous Region:

From the Table I, it can be observed that there is a positive correlation between body weight and body measurement indicators, with the height, chest girth, and chest diameter all showing correlations above 0.5. This indicates that in this dataset, these features may have a strong linear relationship with body weight and can be used to predict it. Therefore, by improving the YOLOv8 algorithm for classifying the body parts of Xinjiang Bactrian camels, the indicators of body length and chest diameter are selected as the main subjects of study in this paper.

TABLE I. PEARSON CORRELATION COEFFICIENTS BETWEEN CAMEL BODY MEASUREMENT FEATURES AND BODY WEIGHT

Feature	Correlation Coefficient
Body Height	0.538
Body Length	0.138
Body Diagonal Length	0.252
Chest Girth	0.793
Chest Diameter	0.594
Teat Diameter	0.256
Teat Length	0.255

B. SKAttention

Different sizes of receptive fields have varying effects on targets of different scales. In the process of classifying different parts of a camel, such as the camel's hump, body, and foot, the appropriate receptive field varies due to the varying sizes of these parts. To address the issue of the camel's body being obscured, we propose the SKAttention mechanism [25], which allows the network to automatically utilize information captured by effective receptive fields for classification. The structure of the Selective Kernel Attention is shown in Fig. 3. And consists of three parts: Split, Fuse, and Select. We input a feature map with dimensions C×H×W, and in the Split part, the input image undergoes convolution operations with 3×3 and 5×5 kernels, resulting in two feature maps, U_1 and U_2. Fuse involves calculating the weights of the two convolution kernels, summing the feature maps element-wise, and then averaging along the H and W dimensions to obtain a one-dimensional vector of size C×1×1. This weight information represents the importance of each channel's information. The formula is as follows:

$$\mathrm{U} = \mathrm{U}_1 + \mathrm{U}_2 \tag{2}$$

In this process, U is generated through global average pooling (AvgPool) to provide channel statistics. A linear transformation is then applied to map the original C-dimensional information to Z-dimensional information. Following this, two linear transformations are used to convert from Z dimensions back to the original C dimensions, thereby completing the extraction of channel dimensions. In the Select part, a softmax operation is performed on the channel dimension to merge the soft attention vectors of each branch, allowing for the selection of multiple branches with different kernel sizes. An SKNet composed of multiple SK units can capture objects of various scales through its neurons.

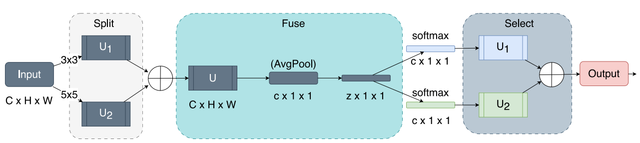

Fig. 3. Selective kernel attention structure diagram.

In the backbone part of YOLOv8, we set the input feature channel number (Channel parameter) to 512, the size list of convolution kernels (kernels parameter) to [1, 3, 5, 7], the dimension reduction ratio (reduction parameter) to 16, the number of convolution groups (group parameter) to 1, and the dimension (L parameter) to 32. Notably, in this process, adopting an adaptive receptive field size allows the model to better adapt to targets of varying scales, thereby enhancing accuracy and efficiency.

C. Asymptotic Feature Pyramid Network AFPN-beta Structure

Inspired by the Asymptotic Feature Pyramid Network (AFPN) [26], we propose a target detection method, AFPN-beta, for detecting various body parts of the Xinjiang Bactrian camel. This method helps to address the issues of target box loss and occlusion among camels, thereby enhancing detection accuracy. The AFPN-beta mainly consists of two key components: the Feature Pyramid Network and the Adaptive Spatial Fusion operation.

The Feature Pyramid Network allows for direct interaction between different layers, preventing the loss of feature information. By extracting features at various scales and integrating them, the model can better capture target features across different scales and semantic levels. This direct interaction helps to improve detection accuracy, especially for targets with complex shapes and varying sizes. The Adaptive Spatial Fusion operation is another crucial component within AFPN-beta. It helps to resolve conflicts of information during the feature fusion process. By dynamically adjusting the fusion weights according to the spatial distribution of features at different layers, the Adaptive Spatial Fusion operation effectively merges features from different levels, reduces information conflicts, and enhances the accuracy of target detection. The specific structure of the AFPN-beta module is illustrated in Fig. 4 [24].

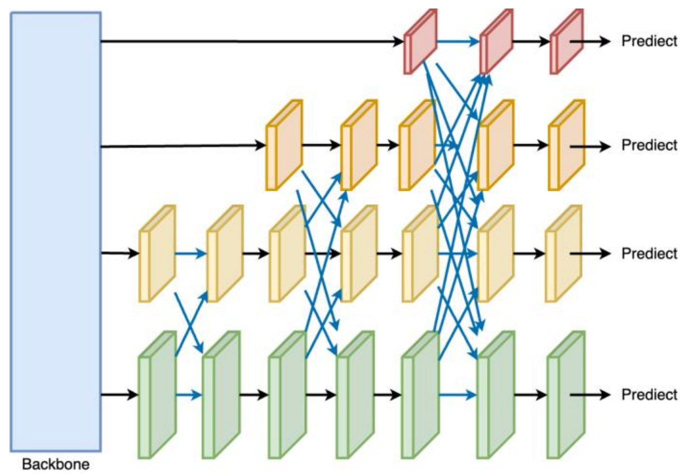

Fig. 4. AFPN-beta module structure diagram.

D. Camel Body Measurement Calculation Algorithm

During the data collection process for Xinjiang Bactrian camels, a set of calibration images using Zhang Zhengyou's calibration board [27] is essential. These images should include a calibration board whose corners have precisely known three-dimensional coordinates. Throughout this process, computer vision techniques such as Harris or Shi-Tomasi corner detection algorithms are utilized to accurately detect these corners on the calibration board. By aligning these detected corners with their established three-dimensional points and employing the camera's projection model, we can accurately estimate the camera's internal parameters—such as focal length and principal point—as well as external parameters, including rotation and translation vectors. Utilizing Zhang Zhengyou's calibration method ensures the precision of the measurement setup, which is crucial for accurate data collection. The calibration process and its results are depicted in Fig. 5, showing the setup used specifically in the data collection of Xinjiang Bactrian camels.

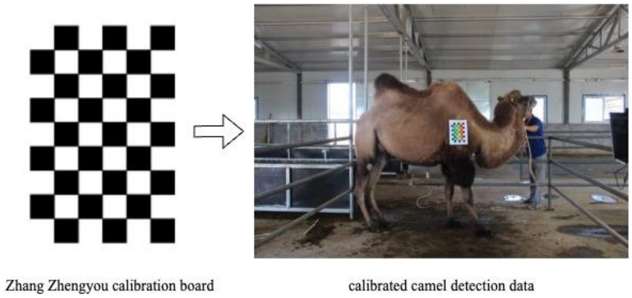

Fig. 5. Apply the Zhang Zhengyou calibration board on the left to the camel data collection on the right.

After the classification is complete, in order to facilitate the extension of the calculation method to three-dimensional space, the Minkowski distance is chosen to calculate the pixel information of various camel parts. The Minkowski distance [28] is a method for measuring the distance between two points in multidimensional space. It is a very common method for measuring the distance between numerical points, with the coordinates of points P and Q assumed to be as follows:

$$P = (x_1, x_2, \ldots, x_n) \text{and } Q = (y_1, y_2, \ldots, y_n) \in R^n \quad (3)$$

Thus, the Minkowski distance is defined as:

$$\left(\sum_{i=1}^{n} |x_i - y_i|^p\right)^{\frac{1}{p}} \quad (4)$$

Once the pixel information is calculated, we can use the similar triangles formula to determine the distance to the object. This process typically requires camera calibration information, including focal length and pixel size parameters. This can be achieved through the internal and external parameters obtained from camera calibration. Below is a basic formula for this conversion:

$$D\_real = \frac{P \times S}{f} \quad (5)$$

Where:

D_real is the actual distance.

f is the focal length of the camera, P= camera_matrix[1,1].

S is the actual distance between the object and the camera.

P is the width of the object in the image.

After completing the calculation of camel body measurements, the mean relative error (MRE) [29] is used to represent the degree of deviation of the measurement error relative to the actual measurement values. The calculation method is shown in the following formula.

$$MRE = \frac{\sum \frac{\text{Actual Value} - \text{Predicted Value}}{\text{Actual Value}}}{n} \quad (6)$$

where, n is the number of samples, Σ represents the sum over all samples, and |x| denotes the absolute value of x.

IV. Experiment

A. Dataset Preparation

Images of Xinjiang Bactrian camels were taken in Ziniquanzi Town, Fukang City, Xinjiang Uygur Autonomous Region, using a Canon 60D camera and an 18-135 lens. Images were collected in December 2020, March to April 2021, and July 2021, capturing 146 Bactrian camels under different lighting conditions, angles, distances, and various occlusions, resulting in 389 valid images. Due to the small dataset size, additional data was collected in September 2023 in Keping County, Aksu Prefecture, Xinjiang Uygur Autonomous Region, amounting to approximately 400 valid images.

A total of 789 images were used to create the dataset for calculating the body measurements of Xinjiang Bactrian camels. The dataset was annotated using the LabelImg software [30].

1) Image selection: Poor quality images were discarded. According to the requirements of the network model, images where key parts of the camel's body were obscured were excluded. Such images were discarded due to overlapping or large occlusions caused by the presence of too many camels in the same frame, which severely affected the training quality of the model. This resulted in 789 valid images.

2) Image annotation: The open-source software LabelImg was used for manual annotation of the targets. SPSS software analysis revealed that four indicators—body size, chest diameter, chest girth, and pipe diameter—are highly correlated with the weight of Bactrian camels. Therefore, the annotated categories include "Camel," "Hump," "Body," "Bottle," and "Foot." The annotations were saved in .txt files with the same names as the images.

3) Data augmentation [31]: The statistical distribution of the sample numbers in the training set showed an imbalance in the number of images of different Xinjiang Bactrian camels. To enhance the model's robustness and adapt to different weather conditions farmers might encounter during camel body detection, techniques such as weather changes, noise addition, and occlusion were randomly applied to the training set images for augmentation.

4) Dataset formatting: The dataset was randomly sampled from images collected at different times and divided into training, validation, and test sets. The training and validation sets together comprise 80% of the dataset, with a 9:1 ratio, while the test set comprises 20%. To verify the model's generalization performance, the collection times of the images in the test set were as different as possible from those in the training set.

B. Training Parameter Settings

The experiments were conducted using an NVIDIA GeForce RTX 3060 GPU, with the operating system being Ubuntu 16.04. The model was built using the Pytorch deep learning framework, with Python version 3.8 and CUDA version 10.2. The improved YOLOv8 model used input images of size 640×640, with an epoch set to 400 and a batch size of 32.

C. Model Evaluation Metrics

The performance of the model was evaluated using precision (P), recall (R), mean average precision (mAP), frame per second (FPS), and memory usage. FPS measures the number of image frames detected per second.

D. Results and Analysis

1) Analysis of body part classification results of xinjiang bactrian camels: In the Xinjiang Bactrian camel dataset, the original training set contained 721 images. Through the application of image augmentation techniques such as weather effect simulation, color jittering, noise addition, and blurring, the training set was expanded to 2518 images for model training. The dataset also included a validation set of 629 images and a test set of 99 images. The model was trained over 400 epochs, and the results of camel body part recognition are shown in Table II.

TABLE II. RESULTS OF COMPARATIVE EXPERIMENTS ON CAMEL PART RECOGNITION

	Epoch	*loss*	*P*	*R*	*mAP 50*	*mAP 50-95*	*FPS*
YOLOv3-tiny	400	0.622	0.745	0.95	0.854	0.69	303.0
YOLOv5n	400	0.648	0.687	0.934	0.804	0.677	285.7
YOLOv5m	400	0.485	0.690	0.943	0.831	0.741	149.2
YOLOv5s	400	0.696	0.730	0.952	0.87	0.717	357.1
YOLOv6n	400	0.569	0.691	0.948	0.833	0.726	357.1
YOLOv6s	400	0.628	0.686	0.943	0.833	0.714	333.3
YOLOv8n	400	0.619	0.691	0.959	0.836	0.719	344.8
YOLOv8s	400	0.611	0.735	0.962	0.879	0.738	303.0
YOLOv10s	400	0.533	0.708	0.923	0.836	0.729	204.1
Camel-YOLOv8	400	0.588	0.745	0.964	0.888	0.764	303.0

From the table above, it is evident that the improved YOLOv8 model performs better in measuring the body dimensions of Xinjiang Bactrian camels compared to YOLOv3, YOLOv5, YOLOv6, YOLOv8, and YOLOv10. The improved YOLOv8 model achieves a precision rate of 74.5%, a recall rate of 96.4%, a mAP50 of 88.88%, a mAP50-90 of 76.4%, and an FPS of 303.03 frames per second. Although the FPS has decreased slightly and the model's memory usage has slightly increased compared to the original YOLOv8 model, other metrics have improved, with the mAP50-95 increasing by 3.7 percentage points. Compared to YOLOv3-tiny, YOLOv5n, YOLOv5s, YOLOv6n, YOLOv6s, YOLOv8n, and YOLOv8s models, the improved model shows an increase in detection accuracy, with the mAP50-95 improving by 7.4, 8.7, 4.7, 3.8, 5.0, 4.5 and 2.6 percentage points, respectively. Additionally, compared to the YOLOv5m and YOLOv10s models, although the bounding box errors have significantly decreased, there has also been a substantial increase in FPS, resulting in a decrease in detection speed.

Therefore, this study successfully enhances the accuracy of body measurement for Xinjiang Bactrian camels while ensuring detection speed, validating the effectiveness of the proposed method. After training for 400 epochs, the training results of the Camel-YOLOv8 model are shown in Fig. 6.

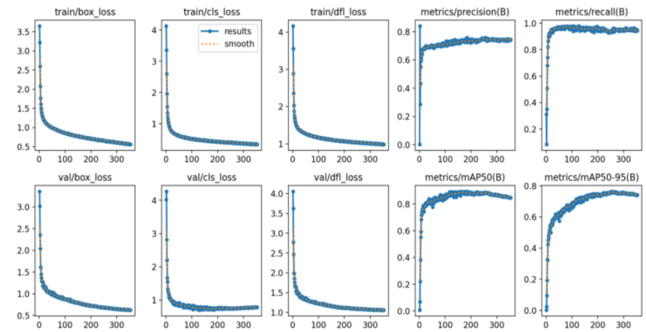

Fig. 6. Training results of the camel-YOLO v8 model.

2) Ablation study: To evaluate the enhancement effects of the proposed improved modules, we designed four ablation experiments. These experiments include the standard YOLOv8 model, YOLOv8 with the SKAttention mechanism, YOLOv8 with the AFPN-beta structure, and a combination of these improvements. Each experimental setup was trained and tested on the same dataset, and consistent evaluation metrics were used to measure changes in model performance. The purpose of these ablation studies is to isolate and quantify the impact of each individual module on the overall performance of the YOLOv8 model. These experiments provide a comprehensive understanding of how each modification contributes to the model's efficacy. The experimental results are presented in Table III.

3) Analysis of Bactrian camel body measurements: The pixel length of the camel's body is determined from the detection bounding box results, and then the actual length of the camel's body is calculated using Formula (6). Partial results of the estimated camel body lengths are shown in Table IV.

Table V presents the error results of this experiment, evaluated using the mean relative error as the metric. The findings indicate that the proposed object detection model demonstrates a high level of accuracy in estimating the pixel-based body measurements of camels.

TABLE III. RESULTS OF ABLATION EXPERIMENTS ON CAMEL PART RECOGNITION

YOLOv8	*SKAttention*	*AFPN-beta*	*Camel-YOLOv8*	*loss*	*P*	*R*	*mAP50*	*mAP50-95*	*FPS*
✓	-	-	-	0.5962	0.720	0.955	0.867	0.734	333.3
✓	✓	-	-	0.5866	0.735	0.952	0.886	0.758	312.5
✓	-	✓	-	0.6111	0.735	0.962	0.879	0.738	303.0
✓	✓	✓	✓	0.5885	0.745	0.964	0.888	0.764	303.0

TABLE IV. PARTIAL RESULTS OF CAMEL BODY LENGTH ESTIMATION

Image Number	*Estimated Camel Body Length (mm)*	*Actual Camel Body Length (mm)*	*Error*	*Estimated Camel Chest Diameter (mm)*	*Actual Camel Chest Diameter (mm)*	*Error*
1	16537	17500	-0.0550	8584.22	9000	-0.0462
2	15817	17100	-0.0750	9890.22	9000	0.0989
3	15045	17000	-0.1150	9849.13	10000	-0.0151
4	16012	17500	-0.0850	11089.13	10000	0.1089
5	15214	16100	-0.0550	9159.29	9300	-0.0151
6	15910	17200	-0.0750	9044.22	9000	0.0049
7	20362	18100	0.1250	9890.22	9000	0.0989
8	18879	17400	0.0850	10837.36	9600	0.1289
9	15970	16900	-0.0550	9847.75	9800	0.0049

TABLE V. RESULTS OF ERRORS IN CAMEL BODY DIMENSION INFORMATION

	MRE
Camel Body Height	-3.39%
Camel Chest Diameter	4.1%

V. CONCLUSION AND FUTURE WORK

A. Conclusion

This research significantly enhances the model's ability to recognize the features of different parts of Bactrian camels, raising the precision, recall, and mean average precision to 74.5%, 96.4%, and 88.8%, respectively. These performance metrics surpass those of the existing YOLOv3, YOLOv5, YOLOv6, YOLOv10, and the original YOLOv8 models. Moreover, the improved model detects at a speed of 303.03 frames per second (fps), demonstrating good real-time performance, and occupies only 14.45MB of memory with fewer parameters, facilitating quick deployment and portability across various devices. Compared to the actual size dimensions of camels, the model's computational results are highly accurate, with an average relative error of -3.39% for height and 4.1% for girth.

In the current context of scaled and refined breeding environments, real-time detection, monitoring, and management of camel populations are becoming increasingly important [32]. Therefore, the outcomes of this study hold significant appeal for camel breeders. By utilizing this model, breeders can more rapidly and precisely assess the growth conditions of camels and calculate their body measurements and weight [33], which is crucial for optimizing selection, improving breeding strategies, and enhancing economic benefits. Additionally, the low error in calculating the body size of Xinjiang Bactrian camels can aid breeders in promptly identifying health issues [34], thereby enabling early intervention measures to reduce the impact of diseases and other health risks on camel breeding.

B. Future Work

Although this study has made significant progress in calculating the body dimensions of camels, there are still some limitations:

1) The model primarily focuses on algorithm optimization for Xinjiang Bactrian camels, and its applicability to other camel breeds or camels from different geographical areas requires further validation.

2) While some technical issues have been addressed through algorithm improvements, the results have not met expectations, particularly in handling occlusions and feature extraction. Therefore, future efforts could explore more effective improvement methods to further enhance the model's accuracy and reliability [35].

Looking ahead, applying this method to other fields will present certain technical challenges. For example, when applied to other animal species, the model may need significant adjustments to accommodate different body features and behavioral patterns. The universality of the model in detecting other livestock or animal body types still needs to be validated and optimized through more experiments. Future research directions include exploring more advanced attention mechanisms and feature fusion techniques to further improve the model's performance in various application scenarios [36][37]. These efforts will help enhance the model's practicality and provide more comprehensive and reliable technical support for intelligent breeding [38].

ACKNOWLEDGMENT

This work was supported by the Natural Science Foundation of Xinjiang Uygur Autonomous Region (No. 2021D01C082).

REFERENCES

[1] Yao Huaibing, Li Na, Liang Xiaorui, Hao Zelin, Yan Hui, Zhao Zhongkai, Liu Ying, Ma Qiang, Wang Yuzhuo, Chen Gangliang, Yang Jie. "Determination and Correlation of Milk Production Traits with Body Size, Body Weight, and Blood Physicochemical Indicators in Camelus bactrianus in Xinjiang." *China Animal Husbandry & Veterinary Medicine*, 2023, Vol. 50, Issue (8): 3210-3220.

[2] N.S.N. Abd Aziz, S.M. Daud, R.A. Dziyauddin, M.Z. Adam and A. Azizan, “A review on computer vision technology for monitoring poultry Farm—Application, hardware, and software,” IEEE access, vol. 9, 2020, pp. 12431-12445.

[3] Yuheng S, Hao Y. "Image segmentation algorithms overview." *arXiv preprint arXiv:1707.02051*, 2017.

[4] Md Golam Morshed, Tangina Sultana, Aftab Alam, and Young-Koo Lee. "Human Action Recognition: A Taxonomy-Based Survey, Updates, and Opportunities." *Sensors*, 2023, Vol. 23, Issue 4, 2182. doi: 10.3390/s23042182.

[5] Suresh Neethirajan. "Artificial Intelligence and Sensor Technologies in Dairy Livestock Export: Charting a Digital Transformation." *Sensors*, 2023, Vol. 23, Issue 16, 7045. doi: 10.3390/s23167045.

[6] Bishop, C. M. "Pattern Recognition and Machine Learning." *Springer*, 2006.

[7] L.O. Chua, “CNN: A vision of complexity,” International Journal of Bifurcation and Chaos, vol. 7, no. 10, 1997, pp. 2219-2425.

[8] B. Koger, A. Deshpande, J.T. Kerby, J.M. Graving, B.R. Costelloe and I.D. Couzin, “Quantifying the movement, behaviour and environmental context of group-living animals using drones and computer vision,” Journal of Animal Ecology, vol. 92, no. 7, 2023, pp. 1357-1371.

[9] M. Madi, Y. Basha, Y. Albadersawi, F. Alenezi, S.A. Mahmoud, et al., “Camel detection and monitoring using image processing and IoT,” Proc. 2023 15th International Conference on Developments in eSystems Engineering (DeSE), IEEE, 2023, pp. 305-308.

[10] M. Hussain, “YOLO-v1 to YOLO-v8, the rise of YOLO and its complementary nature toward digital manufacturing and industrial defect detection,” Machines, vol. 11, no. 7, 2023, pp. 677.

[11] F. Pallottino, R. Steri, P. Menesatti, F. Antonucci, C. Costa, S. Figorilli and G. Catillo, “Comparison between manual and stereovision body traits measurements of Lipizzan horses,” Computers and electronics in agriculture, vol. 118, 2015, pp. 408-413.

[12] Q. Qin, D. Dai, C. Zhang, C. Zhao, Z. Liu, X. Xu, M. Lan, Z. Wang, Y. Zhang and R. Su, “Identification of body size characteristic points based on the Mask R-CNN and correlation with body weight in Ujumqin sheep,” Frontiers in Veterinary Science, vol. 9, 2022, pp. 995724.

[13] K. He, G. Gkioxari, P. Dollár and R. Girshick, “Mask r-cnn,” Proc. Proceedings of the IEEE international conference on computer vision, 2017, pp. 2961-2969.

[14] M. Khojastehkey, M. Yeganehparast, A. Jafari Arvari, N. Asadzadeh and M. Khaki, “Biometric measurement of one-humped camels using machine vision technology,” Journal of Ruminant Research, vol. 7, no. 1, 2019, pp. 19-32.

[15] Zou, Z., Shi, Z., Guo, Y., & Ye, J. "Object Detection in 20 Years: A Survey." *arXiv preprint arXiv:1807.05511*, 2019.

[16] Zhao, Zhong-Qiu, Peng Zheng, Shou-Tao Xu, and Xindong Wu. "Object detection with deep learning: A review." *IEEE Transactions on Neural Networks and Learning Systems*, 2019, Vol. 30, No. 11, pp. 3212-3232.

[17] K. AlNujaidi, G. Alhabib and A. AlOdhieb, “Spot-the-camel: Computer vision for safer roads,” arXiv preprint arXiv:2304.00757, 2023.

[18] K. Wang, H. Guo, Q. Ma, W. Su, L. Chen and D. Zhu, “A portable and automatic Xtion-based measurement system for pig body size,” Computers and electronics in agriculture, vol. 148, 2018, pp. 291-298.

[19] 1. K. Li and G. Teng, “Study on body size measurement method of goat and cattle under different background based on deep learning,” Electronics, vol. 11, no. 7, 2022, pp. 993.

[20] Wang Yusha, Wang Jiaying, Xin Rui, Ke Qiaozhen, Jiang Pengxin, Zhou Tao, Xu Peng. "Application of computer vision in morphological and body weight measurements of large yellow croaker (Larimichthys crocea)." Journal of Fishery Sciences of China, 2023.

[21] M. Ahmad, S. Abbas, A. Fatima, G. F. Issa, T. M. Ghazal, and M. A. Khan, “Deep transfer learning-based animal face identification model empowered with vision-based hybrid approach,” Appl. Sci., vol. 13, no. 2, p. 1178, 2023, DOI: 10.3390/app13021178.

[22] D. Mohanavel and A. M. Ishwarya, “Deep learning and computer vision based warning system for animal disruption in farming environments,” IEEE, 2024, DOI: 10.1109/AIIoT58432.2024.10574724.

[23] Chu Mengyuan, Si Yongsheng, Li Qian, Liu Gang. "Research progress on automatic measurement technology of livestock body size." Transactions of the Chinese Society of Agricultural Engineering, 2022, Vol. 38, No. 13.

[24] Benesty, J., Chen, J., Huang, Y., & Cohen, I. "Pearson Correlation Coefficient." *Noise Reduction in Speech Processing*, Springer, 2009, pp. 1-4.

[25] X. Li, W. Wang, X. Hu and J. Yang, “Selective kernel networks,” Proc. Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, 2019, pp. 510-519.

[26] G. Yang, J. Lei, Z. Zhu, S. Cheng, Z. Feng and R. Liang, “AFPN: Asymptotic feature pyramid network for object detection,” Proc. 2023 IEEE International Conference on Systems, Man, and Cybernetics (SMC), IEEE, 2023, pp. 2184-2189.

[27] Z. Zhang, “A flexible new technique for camera calibration,” IEEE Transactions on pattern analysis and machine intelligence, vol. 22, no. 11, 2000, pp. 1330-1334.

[28] A.A. Thant, S.M. Aye and M. Mandalay, “Euclidean, manhattan and minkowski distance methods for clustering algorithms,” International Journal of Scientific Research in Science, Engineering and Technology, vol. 7, no. 3, 2020, pp. 553-559.

[29] Liu Tong, Teng Guojun, Fu Weidong, Li Zhi. "Extraction algorithms and applications of pig body size measurement points based on computer vision." *Transactions of the Chinese Society of Agricultural Engineering*, 2013, Vol. 29, No. 2, pp. 161-168.

[30] Ali Alameer. "Facial Emotion Recognition Datasets for YOLOv8 Annotation." *University of Salford*, 2023.

[31] Mingle Xu, Sook Yoon, Alvaro Fuentes, and Dong Sun Park. A comprehensive survey of image augmentation techniques for deep learning. Pattern Recognition, page 109347, 2023.

[32] A. A. Samsu Aliar, J. Yesudhasan, M. Alagarsamy, K. Anbalagan, J. Sakkarai, and K. Suriyan, “A comprehensive analysis on IoT based smart farming solutions using machine learning algorithms,” Bull. Electr. Eng. Informatics, vol. 11, no. 3.

[33] C. Tırınk, E. Eyduran, A. Faraz, A. Waheed, N. A. Tauqir, M. Nabeel, M. Tariq, and I. Sheikh, “Use of multivariate adaptive regression splines for prediction of body weight from body measurements in Marecha (Camelus dromedaries) camels in Pakistan,” Trop. Anim. Health Prod., 2021, DOI: 10.1007/s11250-021-02788-y.

[34] M. R. Islam, M. M. Kabir, M. F. Mridha, S. Alfarhood, M. Safran, and D. Che, “Deep learning-based IoT system for remote monitoring and early detection of health issues in real-time,” Sensors, vol. 23, no. 11, p. 5204, 2023.

[35] F. M. J. M. Shamrat, A. Hossain, and T. Roy, “IoT based smart automated agriculture and real time monitoring system,” IEEE.

[36] Y. Hu, X. Deng, Y. Lan, X. Chen, Y. Long, and C. Liu, “Detection of rice pests based on self-attention mechanism and multi-scale feature fusion,” Insects, vol. 14, no. 3, p. 280, 2023.

[37] G. Koc, C. Koc, H. E. Polat, et al., “Artificial intelligence-based camel face identification system for sustainable livestock farming,” Neural Comput. Appl., vol. 36, no. 6, pp. 3107–3124, 2024.

[38] S. Neethirajan, “Recent advances in wearable sensors for animal health management,” Sens. Bio-Sens. Res., vol. 12, pp. 15–29, 2017.