 doi:10.20944/preprints202112.0349.v2

Article

Research on Yak Body Ruler and Weight Measurement Method Based on Deep Learning and Binocular Vision

Wenzhi Wang [1,†], Yuan Zhang[1,†,‡], Jie He[1], Zhanqi Chen[1], Dan Li[1], Chong Ma[1], Yang Ba[2], Qiucuo Baima[2], XiaoQin Li[2], Rende Song[2]

[1] Department of Computer Technology and Applications, Qinghai University, Ningda Road, 810016, Xining, China.
[2] Center for Animal Disease Control and Prevention in Yushu State, 815000, Yushu, China.
† WenZhi Wang and YuAn Zhang have contributed equally to this work
‡ Correspondence address: 2011990029@qhu.edu.cn

Abstract: In order to solve the labor-intensive and time-consuming problem in the process of measuring yak body ruler and weight in yak breeding industry in Qinghai Province, a non-contact method for measuring yak body ruler and weight was proposed in this experiment, and key technologies based on semantic segmentation, binocular ranging and neural network algorithm were studied to boost the development of yak breeding industry in Qinghai Province. Main conclusions: (1) Study yak foreground image extraction, and implement yak foreground image extraction model based on U-net algorithm; select 2263 yak images for experiment, and verify that the accuracy of the model in yak image extraction is over 97%. (2) Develop an algorithm for estimating yak body ruler based on binocular vision, and use the extraction algorithm of yak body ruler related measurement points combined with depth image to estimate yak body ruler. The final test shows that the average estimation error of body height and body oblique length is 2.6%, and the average estimation error of chest depth is 5.94%. (3) Study the yak weight prediction model; select the body height, body oblique length and chest depth obtained by binocular vision to estimate the yak weight; use two algorithms to establish the yak weight prediction model, and verify that the average estimation error of the model for yak weight is 10.78% and 13.01% respectively.

Keywords: yak,semantic segmentation,binocular vision,body ruler,weightestimation

1. Introduction

Pastoral areas in Qinghai Province account for 96% of the total area of the province and the yak in the province account for about 34% of yak population in the world, among which Yushu Prefecture and Guoluo Prefecture have the largest number of yaks[1]. As a unique cattle breed in the alpine region of Qinghai-Tibet Plateau, Yak mainly grows near Qinghai-Tibet Plateau in China, with an altitude of over 3,000 meters and an average temperature below zero. It is the main livestock and economic source of herdsmen in the plateau region [2]. Qinghai Province has the largest number of yaks in the world. After long-term reproduction and growing in natural closed environment in various regions of the province, species resources such as Qinghai plateau yak, Huanhu yak and Xueduo yak with high consistency in body shape and economic characters and stable genetic performance have been formed [3]. In the protection and utilization of yak resources, whether it is the investigation of variety resources, variety selection and matching of yak, sale, calculation of ration and dosage, etc., it is necessary to take the weight of yak as the basis [4]. Because yak is big and strong, manual measurement needs the cooperation of multiple people to complete the statistics of the whole set of body ruler data of a yak [5].

In 2003, Teng Guanghui and used visual methods to measure the weight and body ruler of pigs [6]. In 2006, Huang Junran of Hebei Agricultural University completed

 doi:10.20944/preprints202112.0349.v2

the measurement of dairy cow's body ruler by using reference object calibration, and designed and accomplished the management system of dairy cow's image and information [7]. The images of lactating cows were used to evaluate their physical condition in 2008. When the dairy cows passed through the weighing station, the images of the cows were taken, and the USBCS and UKBCS physical condition evaluation system were used to evaluate their physical condition [8]. In 2011, Tasdemir used digital image analysis method to measure the body ruler data of Holstein cows, and established the weight prediction model of Holstein cows based on fuzzy rules through regression analysis of the measured data[9]. In 2013, in order to study the comparison of dairy cow contour segmentation algorithms in natural image background, T.Van Hertem and others evaluated five different target segmentation algorithms of background images in static and dynamic background, respectively. The results show that the extraction effect of dairy cow body contour in static background is obviously better than that in dynamic background[10]. In 2014, Zhang Wen of the Inner Mongolia University of Science & Technology designed and implemented a sheep physical sign monitoring system using the algorithm in OpenCV visual library[11]. In 2015, Yukako Kuzuhara et al. used ASUS Xtion Pro sensor to measure dairy cow's back posture. By measuring and recording six characteristic positions of the dairy cow, the linear regression algorithm was used to evaluate the dairy cow's physical condition. This verified that the three-dimensional camera system could cow body measurement and performance analysis[12]. In 2016, Doeschl et al. recorded the growth process of pigs by using a visual image analysis system based on a single camera, and established the relationship between pig body ruler parameters and time [13]. In 2021, Sun Zijie used machine vision to estimate yak weight, and adopted traditional image processing methods to obtain yak foreground images[14].

In this study, mobile devices are used to acquire images of yaks, and the parameters of yaks are accurately obtained under non-contact conditions. The research contents are: 1) The acquired binocular images of yaks are trained by U-net network to get foreground images of yaks. 2) Marking the measurement points in combination with the segmented yak foreground image; finding the corresponding position in the yak depth map formed by the binocular image to obtain the position information of the yak body ruler measurement points; acquiring the yak body ruler data combining the position information of the key measurement points. 3) Finally, the yak weight was estimated. The flow chart of the measurement method in this paper is shown in Figure 1. This study can save time and labor costs, guide herders to manage yak reasonably, and provide a guarantee for the sustainable development of alpine pastoral areas.

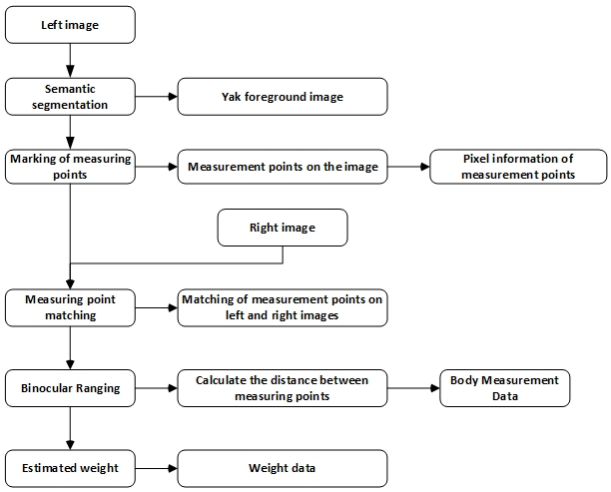

Figure 1. Flow of the Measurement Method.

 doi:10.20944/preprints202112.0349.v2

2. Material and Method

2.1. Capture of Yak Image

The yak images used in this experiment were from Qumalai County, Yushu Tibetan Autonomous Prefecture, Qinghai Province. 2,263 binocular images of 200 yaks were captured, and the body ruler data was obtained through manual measurement. The weight data of 40 yaks were obtained through electronic weighing. The experimental image is shown in Figure 2. Part of the server platform of the test semantic segmentation is configured with Intel Xeon E5-2603 as the CPU, GeForce GTX1080Ti as the graphics card, and the binocular and weight estimation are completed on this machine with the Core i5-9400 as the CPU.

Figure 2. Experimental Image.

2.2. Extraction of Yak Foreground Targets

At present, many scholars have applied semantic segmentation algorithm to animal image segmentation[15]. For non-contact measurement of weight and body ruler of yak, first, the collected yak images should be separated from the background, and then the weight and body ruler are estimated by combining the binocular vision. We use U-net algorithm to observe the overall situation first, roughly find the target area[16], and then consider the more detailed information for further judgment. This can make the segmentation results more accurate[17]. Therefore, the U-net algorithm is used for yak image segmentation.

The U-net algorithm was adopted for semantic segmentation of the yak's foreground and background, with two categories[18]. The image was manually segmented by Labelme, and the foreground and background of yak's image were separated to obtain a json file in coco dataset format. The image shown in Figure 3 was obtained after program processing. The data set was enhanced by increasing the sample size to 3,409 to improve the generalization ability. The obtained data set were randomly divided into the training set, the test set, and the verification set according to the ratio of 7:2:1, and the data set was input into the network of U-net[19]. The pre-training model of the VOCdevkit data set was adopted, and the iteration was performed for 200 times. The random gradient descent method was used, and the learning rate would have been 0.0001, Batch_size to be 8, and base_size to be 521×521, then the final training model is obtained[20].

U-net, one of the earlier semantic segmentation algorithms employing full convolutional network, uses a u-shaped structure including a compression path and an expansion path. The U-net network mainly includes up-sampling part and down-sampling part to increase the receptive field and improve the most obvious characteristics of the image through maximum pooling[21,22]. In this way, the down-sampling process is completed, and the image is mapped from small resolution to large resolution by up-convolution to restore the image size and complete up-sampling. The entire network uses convolution and rectified linear function to complete the sampling of the input image and outputs a feature map. Among them, convolution is mainly used to extract local features of images, while rectified linear function is mainly used to keep relevant features and remove irrelevant features[23]. Introducing skip-connection in U-net network can lead out shallow convolution features. In the process of up-sampling at each level, the U-net network fuses the feature map of the corresponding positions of encoders on the channel[24], thus ensuring that the finally restored images fuse more bottom features and features

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 9 March 2022 doi:10.20944/preprints202112.0349.v2

of different scales, keeping more world information in segmentation and improving segmentation accuracy.

Figure 3. Original Image and Manually Segmented Image.

2.3. Marking of Measurement Points of Yak body ruler

After segmenting the foreground image of the yak by the U-net network, the measurement points of the body ruler of the yak on the foreground image are marked. Accurately marked measurement points can make people precisely find the corresponding positions of the measurement points in space on the depth image of the yak combining the binocular image, thus further exactly calculating the body ruler data of the yak[25,26]. In this experiment, the yak body height, body oblique length and chest depth were selected to calculate the yak weight, so the measurement points related to the yak body height, body oblique length and chest depth should be marked on the yak foreground image. Table 1 illustrates the positions of the measurement points in the image, and the specific positions of the measurement points are shown in Figure 4.

Table 1. Description of Corresponding Positions of Measurement Points.

body ruler	Description	Position on the Image
Body Height	The distance from the withers to the ground	A1-A2
Body Oblique Length	Distance between shoulder and ischial end	B1-B2
Chest Depth	Distance from back end of scapula to abdomen	C1-C2

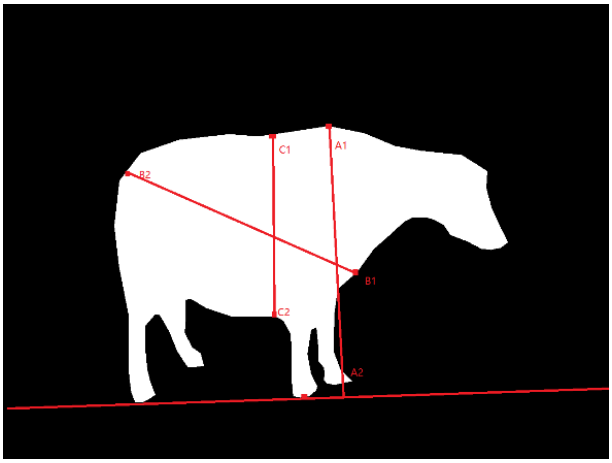

Figure 4. Original Image and Manually Segmented Image.

According to Figure 4, it is necessary to mark the measurement points of yak body ruler in the foreground image of yak segmented by U-net network. First, it is essential to get the yak contour from the foreground image, as shown in Figure 5. Yak contour is the edge of the segmented yak foreground image, and a reflection of local color abrupt change to the image. Calculate the average value of color values in 8 directions for each

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 9 March 2022 doi:10.20944/preprints202112.0349.v2

pixel of the foreground image, and set 254 and 136 as the upper and lower thresholds. The pixels in this range are the contour of yak foreground image.

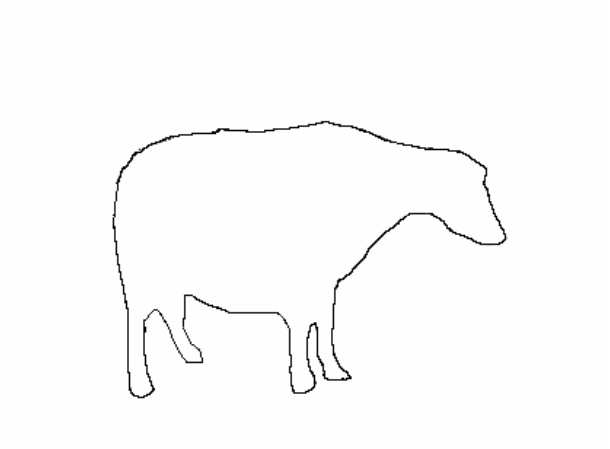

Figure 5. Contour.

After obtaining the yak contour, it is necessary to find out the six measurement points related to the yak body ruler, as shown in Figure 4, and calculate the geometric center of the extracted yak foreground image.

2.4. Use Binocular Vision to Measure Yak body ruler

After the segmentation of left and right images and the marking of measurement points, we enter the binocular ranging stage. As an important branch of computer vision, binocular vision uses two cameras in different positions to shoot the same scene[27]. As shown in Figure 6, the binocular camera used in this experiment can capture two images of 1280×960, and obtain the 3D coordinates of the spatial point by calculating the parallax of the spatial point in the two images[28,29]. The principle of binocular imaging is shown in Figure 7.

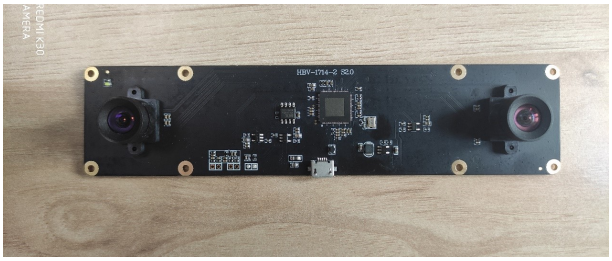

Figure 6. Binocular Camera.

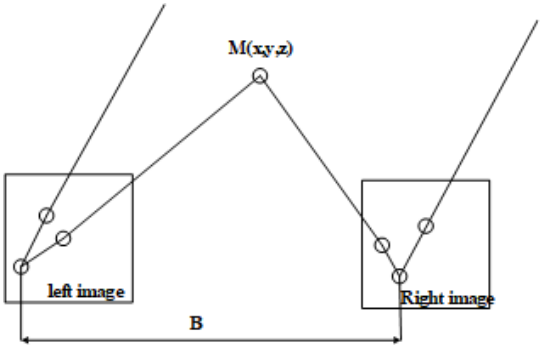

Figure 7. Principle of Binocular Imaging.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 9 March 2022 doi:10.20944/preprints202112.0349.v2

As shown in Figure 7, B is the distance between the projection centers of the two cameras. Let's assume a measurement point $M(x,y,z)$ in space, and the image points of M in binocular cameras are respectively $M\_left = (x\_left, y\_left)$ and $M\_right = (x\_right, y\_right)$. Because the two cameras are in the same plane, so the coordinate y of the measurement point images on the left and right cameras are same. That is, $y = y\_left = y\_right$, and the three-dimensional coordinates of M can be obtained from the triangular geometric relationship: $x = (Bx\_left)/d$,$y = By/d$,$z = Bf/d$, where f is the focal length of the camera $d = x\_left - x\_right$ [30,31]. Therefore, any point on the image of the left camera can determine the 3D coordinates of the point as long as the corresponding matching point can be found on the image of right camera[32].

Before binocular ranging, it is necessary to calibrate the camera to obtain intrinsic and extrinsic parameters[33]. We select Zhang Zhengyou chessboard calibration method[34], and obtain the focal length f, baseline distance B of the binocular camera, and the translation vector T and rotational vector R between two cameras through calibration. Table 2 shows the data obtained by binocular calibration.

Table 2. Calibration Data.

	Left Camera	**Right Camera**
Intrinsic matrix	$\begin{bmatrix} 1465.26 & -1.051 & 692.14 \\ 0 & 1465.93 & 444.83 \\ 0 & 0 & 1 \end{bmatrix}$	$\begin{bmatrix} 1692.14 & -1.42 & 669.93 \\ 0 & 1464.19 & 455.80 \\ 0 & 0 & 1 \end{bmatrix}$
Distortion Vector	$\begin{bmatrix} 0.25 & -1.22 & -0.007 & 0.004 & 0 \end{bmatrix}$	$\begin{bmatrix} 0.199 & -1.097 & -0.001 & 0.002 & 0 \end{bmatrix}$
Intrinsic matrix	$\begin{bmatrix} -116.355 & -1.774 & -5.994 \end{bmatrix}$	
Distortion Vector	$\begin{bmatrix} 0.999999 & 0.00038 & -9.098e\text{-}5 \\ -0.00038 & 0.99998 & -0.00524 \\ 8.896e\text{-}5 & 0.00524 & 0.99998 \end{bmatrix}$	

According to the principle of three-dimensional coordinates of a point in the binocular measured image, and the data obtained by camera calibration, combined with the measurement points found in the foreground image, the corresponding positions of the measurement points can be found in the depth image, thus obtaining the three-dimensional coordinates and yak body ruler parameters.

2.5. Yak Weight Estimation

There is a high correlation between the weight and body ruler of yak[35]. In this study, the body height, body oblique length and chest depth of yak were selected for modeling by using method such as linear regression[36] and support vector regression[37]. Six yak weight prediction models were established based on the old and new data. The correlation between body ruler and weight was analyzed by using 3581 data. The results are shown in Table 3. It can be seen that the body ruler data has a high correlation with weight. In the data set, W, tg, tx, xw show the weight, the body height, the body oblique length, and the chest circumference respectively.

Table 3. Correlation Coefficient Between body ruler and Weight of Yak.

	W	tg	tx	xw
W	1.000000	0.749930	0.653193	0.693756
tg	0.749930	1.000000	0.636013	0.649566
tx	0.653193	0.636013	1.000000	0.543669
xw	0.693756	0.649566	0.543669	1.000000

In the linear regression algorithm model, each data has n features, and each feature corresponds to its own weight value, and the product with the weight plus an offset value is the linear regression model[38]. The formula is:

$$y = w_1 * x_1 + w_2 * x_2 + ... + w_n * x_n + b \tag{1}$$

If $w_0 = b$,$x_0 = 1$, then you can get:

$$y = w_0 * x_0 + w_1 * x_1 + w_2 * x_2 + ... + w_n * x_n \tag{2}$$

Now there are m samples to get the matrix represented as:

$$X = \begin{bmatrix} 1 & x_1^1 & x_1^2 & ... & x_1^n \\ 1 & x_2^1 & x_2^2 & ... & x_2^n \\ ... & ... & ... & ... & ... \\ 1 & x_m^1 & x_m^2 & ... & x_m^n \end{bmatrix} Y = \begin{bmatrix} y_1 \\ y_2 \\ ... \\ y_m \end{bmatrix} \tag{3}$$

Among them, the weight w is represented as:

$$w = [w_0 \quad w_1 \quad w_2 \quad ... \quad w_n] \tag{4}$$

The model function of support vector regression model is also a linear function, but it differs from the linear regression model in the principle of calculating loss and the aim function and optimization algorithm. A "marginal zone" is created on both sides of the linear function of support vector regression, and no loss is calculated for all samples falling into the marginal zone[39]. Only those outside the marginal zone are included in the loss function. Then, the model is optimized by minimizing the width and total loss of the marginal zone.

3. Experimental Results and Analysis

3.1. Evaluation of Foreground Extraction Results

The evaluation indexes of semantic segmentation mainly include three indexes: execution time, memory usage, and accuracy. Considering this experiment, four kinds of semantic segmentation accuracy evaluation indexes are used to evaluate the segmentation model[40,41]. They are PA (pixel accuracy), CPA (category pixel accuracy), MPA (mean pixel accuracy) and MIoU (mean intersection over union).

Assuming that there must be $k + 1$ categories (including k target categories and 1 background category), the pixels with i as p_ii identification category are predicted as the total number of i; the pixels with i as p_{ij} identification category are predicted as the total number of j; and the pixels with j as p_{ji} identification category are predicted as the total number of i.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 9 March 2022 doi:10.20944/preprints202112.0349.v2

PA is the pixel accuracy, that is the ratio of the number of correctly classified pixels to the number of all pixels is represented as:

$$PA = \frac{\sum_{i=0}^{k} p_{ii}}{\sum_{i=0}^{k} \sum_{j=0}^{k} p_{ij}} \tag{5}$$

CPA is the category pixel accuracy, that is the accuracy of pixels that really belong to category i in category i prediction is represented as:

$$CPA_i = \frac{\sum_{i=0}^{k} p_{ii}}{\sum_{i=0}^{k} p_{ii} + \sum_{j=0}^{k} p_{ij}} \tag{6}$$

MPA is the mean pixel accuracy, that is the average value of the ratio between the number of correctly classified pixels in each category and the number of all pixels in that category is represented as:

$$MPA = \frac{1}{k} \sum_{i=1}^{k} \frac{p_{ii}}{\sum_{i=0}^{k} p_{ij}} \tag{7}$$

MIoU is the mean intersection over union, that is the average value of IoU of each category (the ratio of intersection and union of the predicted results and true values of a certain category of the model) is represented as:

$$MIoU = \frac{1}{k+1} \sum_{i=1}^{k} \frac{p_{ii}}{\sum_{i=0}^{k} p_{ij} + \sum_{i=0}^{k} p_{ji} - p_{ii}} \tag{8}$$

In this study, the training was repeated many times, and the results of the model prediction accuracy obtained are shown in Table 4.

Table 4. Evaluation of Model Prediction Accuracy.

No.	PA	CPA_0	CPA_1	MPA	MIoU
1	0.9831	0.9863	0.9714	0.9789	0.9506
2	0.9839	0.9868	0.9728	0.9798	0.9535
3	0.9834	0.9884	0.9658	0.9771	0.9519
4	0.9831	0.9849	0.9761	0.9805	0.9510
5	0.9836	0.9857	0.9753	0.9805	0.9520
6	0.9833	0.9856	0.9746	0.9801	0.9516
7	0.9825	0.9854	0.9722	0.9788	0.9490
8	0.9842	0.9871	0.9731	0.9801	0.9537
9	0.9831	0.9854	0.9753	0.9803	0.9514
10	0.9833	0.9873	0.9696	0.9784	0.9515
Mean Value	0.9833	0.9862	0.9726	0.9794	0.9516

From Table 6, it can be seen that after 10 times of repeated training, the average PA of the model is 0.9833. For CPA, the background accuracy and the foreground accuracy are 0.9862 and 0.9722, respectively. The MPA is 0.9794, and the MIoU is 0.9516. This shows that the model has good stability and generalization ability.

By comprehensively comparing the results of 10 times training, the accuracy of the eighth training is higher. Therefore, the U-net network structure of the eighth training is stored for subsequent direct call. We can use the trained model for yak image segmentation. The results are shown in Figure 8. According to the segmentation results of yak image, it can be seen that the segmentation of yak foreground image based on U-net network built in this experiment has better results, and it also has better segmentation performance on complex background images.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 9 March 2022 doi:10.20944/preprints202112.0349.v2

Figure 8. Yak Image Segmentation Result. (**a**) Original Image. (**b**) Segmented Image.

3.2. Evaluation of Measurement Results of Yak body ruler

Party binocular images need to be selected for the experiment. We use U-net network to get the left and right yak foreground images, and find the measurement points related to body ruler on the foreground images, as shown in Figure 9. It is the marked measurement points related to body ruler. Then we get the pixel information related to the measurement points, and obtain the depth map in combination with the binocular images, and find the positions corresponding to the measurement points in the depth map. We acquire the yak body ruler data, including body height, body oblique length and chest depth. Table 5 is the result statistics of body ruler measurement data, including data for 90 yaks.

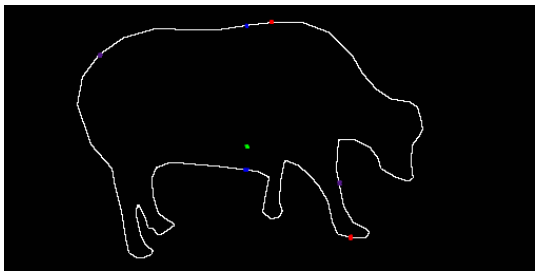

Figure 9. Measurement Point.

Table 5. Result Statistics of body ruler Measurement.

	Max error/cm	Min Error/cm	Ave Error/cm
Body Height	4.699	1.121	2.594
Body Oblique Length	5.329	1.147	2.666
Chest Depth	11.996	1.486	5.94

The data of 10 yaks were randomly selected to show the results, as shown in Table 6.

 doi:10.20944/preprints202112.0349.v2

Table 6. Part of Yak Data.

Body Height			Body Oblique Length			Chest Depth		
Manual	Binocular	Error	Manual	Binocular	Error	Manual	Binocular	Error
104	106.045	1.97%	109	106.92	1.91%	40	42.91	7.80%
104	108.699	4.52%	108	110.055	1.90%	40	42.76	6.57%
104	107.252	3.13%	110	108.406	1.45%	39	41.77	7.50%
106	109.249	3.07%	114	119.329	4.68%	43	41.56	2.62%
105	103.872	1.07%	109	106.286	2.49%	42	46.14	8.94%
104	106.311	2.22%	110	113.123	2.84%	43	46.5	8.95%
105	103.676	1.26%	115	112.593	2.09%	46	47.62	3.83%
106	109.622	3.42%	113	115.007	1.78%	44	45.91	4.45%
105	108.906	3.72%	112	116.446	3.97%	45	43.87	2.31%
104	108.501	4.33%	109	104.463	4.16%	45	45.49	1.31%

Yak body ruler measurement based on binocular vision. It combines the foreground image obtained by U-net network and finds the relevant measurement points on the foreground image, then matches the features on the binocular image and seeks the measurement points to obtain yak body ruler data. According to the test results, the measurement results of this method are not much different from the manually measured yak body ruler data, but the overall cost will be much less than that of manual measurement. It can save the time and labor cost in measuring yak body ruler data.

3.3. Evaluation of Yak Weight Estimation Results

3580 yaks body ruler and weight data are selected to fit the model which predicting weight with yak body ruler. Among them, the data for 200 yaks is new. Linear regression algorithm and support vector regression algorithm are adopted to get the yak weight prediction model. The overall performance of the four yak weight prediction models on the test data are as shown in Figure 10.

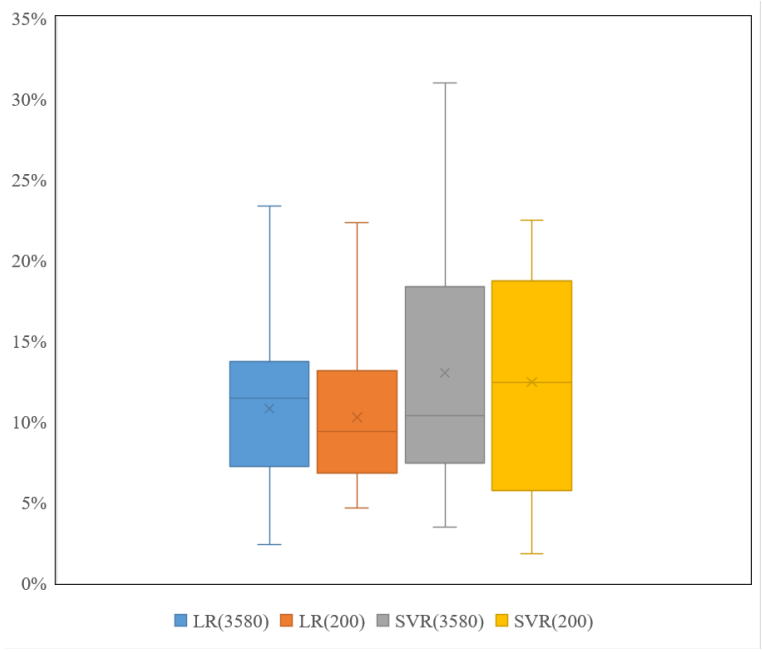

Figure 10. Yak weight estimates the overall performance of the model.

The yak body ruler data obtained by binocular measurement is used to estimate the yak weight. For example, Table 7 shows the average error and fitting equation of the four models based on 90 groups of test data.

Table 7. Result Statistics of body ruler Measurement.

Method	Equation	AVG /%
LR (3580)	$W = 3.7775 * BH + 1.0258 * BOL + 0.8317 * CD - 502.0926$	10.78
SVR (3580)		13.01
LR (200)	$W = 3.8724 * BH + 2.0787 * BOL + 0.1364 * CD - 478.207$	10.26
SVR (200)		12.44

4. Conclusion and discussion

In this paper, the images taken by binocular cameras are used as experimental materials to estimate yak's body ruler and weight under non-contact conditions. The experimental results show: (1) The segmentation accuracy of yak image foreground extraction algorithm based on U-net network is over 97%. (2) Based on binocular vision, the average measurement error of yak body height and body oblique length is 2.6%, and the average measurement error of yak chest depth is 7%. (3) We estimate the yak weight according to yak body ruler data based on semantic segmentation and obtained by binocular vision, the average errors of the four prediction models are 10.78%, 13.01%, 10.26% and 12.44%, respectively. Based on the above results, the non-contact measurement method of yak body ruler and weight proposed in this experiment have high accuracy, and can effectively solve the time and labor consuming problem in the process of yak body ruler and weight measurement under the condition of ensuring the accuracy of binocular images.

Author Contributions: Conceptualization, W.W. and Y.Z.; methodology, W.W.; software, W.W.; vali-dation, S.Z., G.L. and W.Z.; formal analysis, W.W.; investigation, Y.B.,Q.B.,and X.L.; re-sources, J.H.,Z.C.,C.M.and D.L.; data curation,W.W.; writing-original draft preparation, W.W.; writing—review and editing, W.W.; visualization,W.W.; supervision, R.S.; funding acquisition, Y.Z. All authors have read and agreed to the published version of the manuscript.

Funding: This work is supported by Science and Technology Project in Qinghai Province (No: 2020-QY-218); Supported by China Agriculture Research System of MOF and MARA(CARS-37); And received the support of the "High-end Innovative Talents Thousand Talents Program" in Qinghai Province.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Limited data available on request due to the large size of the data.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

LR	Linear Regression
SVR	Support Vector Machine Regression
W	Weight
BH	Body Height
BOL	Body Oblique Length
CD	Chest Depth
AVG	Average Value

References

1. A. Cy, B. Gt, A. Qf, A. Sw, A. Zw, A. Sc, and A. Fh, "Behavioral patterns of yaks (bos grunniens) grazing on alpine shrub meadows of the qinghai-tibetan plateau," *Applied Animal Behaviour Science*, vol. 234, 2020.
2. D. U. Zi-Yin, Y. J. Cai, X. D. Wang, and B. Zhang, "Research progress on yak grazing behavior and its influence on the soil properties of alpine grassland," *Acta Prataculturae Sinica*, 2019.

3. Qi, Yan, Luming, Ding, Haiyan, Wei, Xianju, Wang, and Cuixia, "Body weight estimation of yaks using body measurements from image analysis," *Measurement*, 2019.
4. C. Huang, F. Ge, W. Ren, Y. Zhang, and C. Liang, "Copy number variation of the hpgds gene in the ashidan yak and its associations with growth traits," *Gene*, vol. 772, no. 1, p. 145382, 2020.
5. Y. Qiao, H. Kong, C. Clark, S. Lomax, and S. Sukkarieh, "Intelligent perception for cattle monitoring: A review for cattle identification, body condition score evaluation, and weight estimation," *Computers and Electronics in Agriculture*, vol. 185, 2021.
6. Li. C, G. Teng , C. Zhao, "Using computer vision technology to realize non-destructive monitoring of greenhouse plant growth," *Transactions of the Chinese Society of Agricultural Engineering*, vol. 03,pp.140-143, 2003.
7. H. J, "Research on Web-based Cow Image Recognition and Image Information Management System," *Hebei Agricultural University*,2006.
8. J. M. Bewley, A. M. Peacock, O. Lewis, R. E. Boyce, D. J. Roberts, M. P. Coffey, S. J. Kenyon, and M. M. Schutz, "Potential for estimation of body condition scores in dairy cattle from digital images," *Journal of Dairy Science*, vol. 91, no. 9, pp. 3439–3453, 2008.
9. S. Tasdemir, A. Ürkmez, and S. Inal, "A fuzzy rule-based system for predicting the live weight of holstein cows whose body dimensions were determined by image analysis," *Turkish Journal of Electrical Engineering & Computer Sciences*, vol. 19, no. 4, pp. 689–703, 2011.
10. T. V. Hertem, V. Alchanatis, A. Antler, E. Maltz, I. Halachmi, A. Schlageter-Tello, C. Lokhorst, S. Viazzi, C. Romanini, and A. Pluk, "Comparison of segmentation algorithms for cow contour extraction from natural barn background in side view images," *Computers & Electronics in Agriculture*, vol. 91, pp. 65–74, 2013.
11. W. Zhang, "Research on livestock physical sign measurement system based on machine vision," *Inner Mongolia University of Science and Technology*, 2014.
12. Y. Kuzuhara, K. Kawamura, R. Yoshitoshi, T. Tamaki, S. Sugai, M. Ikegami, Y. Kurokawa, T. Obitsu, M. Okita, and T. Sugino, "A preliminarily study for predicting body weight and milk properties in lactating holstein cows using a three-dimensional camera system," *Computers and Electronics in Agriculture*, vol. 111, pp. 186–193, 2015.
13. A. B. Doeschl-Wilson, C. T. Whittemore, P. W. Knap, and C. P. Schofield, "Using visual image analysis to describe pig growth in terms of size and shape," *Animal Science An International Journal of Fundamental & Applied Research*, vol. 79, no. 03, pp. 415–427, 2016.
14. Zhang. Y A, Sun. Z J, Zhang. C. Whittemore, P. W. Knap, and C. P. Schofield, "Body weight estimation of yak based on cloud edge computing," *J Wireless Com Network 2021* .vol. 6,2021.
15. S. Zhao, G. Hao, Y. Zhang, and S. Wang, "A real-time semantic segmentation method of sheep carcass images based on icnet," *Journal of Robotics*, vol. 2021, no. 2, pp. 1–12, 2021.
16. O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, 2015.
17. G. Martinez-Soltero, A. Y. Alanis, N. Arana-Daniel, and C. Lopez-Franco, "Semantic segmentation for aerial mapping," 2020.
18. D. Xu, X. Zhou, X. Niu, and J. Wang, "Automatic segmentation of low-grade glioma in mri image based on unet++ model," *Journal of Physics: Conference Series*, vol. 1693, no. 1, p. 012135 (7pp), 2020.
19. I. C. Saidu and L. Csató, "Active learning with bayesian unet for efficient semantic image segmentation," *Journal of Imaging*, vol. 7, no. 2, p. 37, 2021.
20. H. Jiang and Y. E. Xining, "An improved skin disease image segmentation algorithm based on i-unet network," *Modern Electronics Technique*, 2019.
21. Y. A. Ayalew, A. Kinde, and M. A. Mohammed, "Modified u-net for liver cancer segmentation from computed tomography images with a new class balancing method," *BMC Biomedical Engineering*, vol. 3, no. 1, 2021.
22. S. T. Tran, C. H. Cheng, T. T. Nguyen, M. H. Le, and D. G. Liu, "Tmd-unet: Triple-unet with multi-scale input features and dense skip connection for medical image segmentation," *Healthcare*, vol. 9, no. 1, p. 54, 2021.
23. Y. F. Zhang, J. Zheng, L. Li, N. Liu, and X. He, "Rethinking feature aggregation for deep rgb-d salient object detection," *Neurocomputing*, vol. 423, no. 4, pp. 463–473, 2021.
24. A. E. Maxwell, M. S. Bester, L. A. Guillen, C. A. Ramezan, and J. L. Pyron, "Semantic segmentation deep learning for extracting surface mine extents from historic topographic maps," *Remote Sensing*, vol. 12, no. 24, p. 4145, 2020.
25. P. Azad, *Visual Perception for Manipulation and Imitation in Humanoid Robots*. Visual Perception for Manipulation and Imitation in Humanoid Robots, 2009.
26. L. Li, B. Li, Z. Wu, D. Gao, and X. Mu, "Automatic placement of annotation in area feature by map spatial geometry information measurement," *Proceedings of SPIE - The International Society for Optical Engineering*, vol. 6751, 2007.
27. L. Zhang, C. Li, Y. Fan, X. Zhang, and J. Zhao, "Physician-friendly tool center point calibration method for robot-assisted puncture surgery," *Sensors*, vol. 21, no. 2, p. 366, 2021.
28. Z. C. Qiu and Z. Q. Huang, "A shape reconstruction and visualization method for a flexible hinged plate using binocular vision," *Mechanical Systems and Signal Processing*, vol. 158.
29. D. Ruan, W. Zhang, and D. Qian, "Feature-based autonomous target recognition and grasping of industrial robots," *Personal and Ubiquitous Computing*, no. 1, pp. 1–13, 2021.
30. F. Jin and Y. H. Fan, "Digital image stabilization with motion estimation based on binocular ranging and sensor," in *Chinese Control Conference*, 2014.

31. T. Shen, W. Liu, and J. Wang, "Distance measurement system based on binocular stereo vision," *Electronic Measurement Technology*, 2015.
32. C. Shi, J. Zhang, and G. Teng, "Mobile measuring system based on labview for pig body components estimation in a large-scale farm," *Computers and Electronics in Agriculture*, vol. 156, pp. 399–405, 2019.
33. Dingfei, Jin, Yue, and Yang, "Using distortion correction to improve the precision of camera calibration," *Optical Review*, vol. 26, no. 2, p. 269–277, 2019.
34. A. Zisserman, "Book review : Epipolar geometry in stereo, motion and object recognition—a unified approach by gang xu and zhengyou zhang published by kluwer academic publishers group; 1996; 313 pages; us$ 160," *International Journal of Robotics Research*, vol. 17, no. 8, pp. 903–904, 1998.
35. A. J. da Silva Cardoso, C. A. L. D. Oliveira, E. C. Campos, R. P. Ribeiro, and F. F. E. Silva, "Estimation of genetic parameters for body areas in nile tilapia measured by digital image analysis," *Journal of Animal Breeding and Genetics*, 2021.
36. M. Lira, E. Silva, J. M. B. Alves, and G. Veras, "Estimate of wind resources in the coast of ceará using the linear regression theory," *Renewable & Sustainable Energy Reviews*, vol. 39, no. nov., pp. 509–529, 2014.
37. T. K. Lin, T. H. Yang, P. H. Wang, R. C. Zeng, and K. C. Chang, "Prediction of smooth hysteretic model parameters using support vector regression," *Multiscal*
38. L. I. Dahu,"Predicting short-term traffic flow in urban based on multivariate linear regression model," *ournal of Intelligent and Fuzzy Systems*.vol 39, pp. 1–11, 2020.
39. P. Anand, R. Rastogi, S. Chandra, A class of new support vector regression models, Applied Soft Computing (2020) 106446.
40. M. Shahzad, A. I. Umar, M. A. Khan, S. H. Shirazi, Z. Khan, W. Yousaf, Robust method for semantic segmentation of whole-slide blood cell microscopic image, Computational and Mathematical Methods in Medicine 2020 (2020) 1–13.
41. F. Lateef, Y. Ruichek, Survey on semantic segmentation using deep learning techniques, Neurocomputing 338 (2019) 321–348