

Automatic body condition scoring for dairy cows based on efficient net and convex hull features of point clouds



Kaixuan Zhao^{a,b}, Meng Zhang^{a,b}, Weizheng Shen^c, Xiaohang Liu^a, Jiangtao Ji^{a,b,*}, Baisheng Dai^{c,*}, Ruihong Zhang^a

^a College of Agricultural Equipment Engineering, Henan University of Science and Technology, Luoyang 471023, China

^b Henan International Joint Laboratory of Intelligent Agricultural Equipment Technology, Henan University of Science and Technology, Luoyang 471023, China

^c College of Electrical Engineering and Information, Northeast Agricultural University, Harbin 150030, China

ARTICLE INFO

Keywords:

Body condition score
Convex hull feature
EfficientNet network
WOA
Two-level model

ABSTRACT

Body condition score (BCS) is an important indicator to evaluate the nutritional and health status of perinatal dairy cows. The insufficient accuracy of the automatic body condition score is the core bottleneck restricting the popularization and application of this technology. To further improve the accuracy of automatic scoring of cow body condition, a 3D structure feature map based on convex hull distance of point clouds was constructed and used as the input of a two-level model based on EfficientNet network, which realized the improvement of recognition accuracy within the error of 0.25 in the automatic scoring of cow body condition. First, the depth image of the cow's back was preprocessed to obtain the point cloud from the hook bone to the needle bone region containing the main body condition information. Second, the point cloud was voxelized and convex hulled, and the distance between each peripheral voxel and the nearest convex hull surface was calculated and projected to the X-Y plane to obtain the structural feature map. Then, a two-level classification model based on the EfficientNet network was constructed, and the whale optimization algorithm (WOA) was used to optimize the network scaling coefficient in the model. Finally, 5119 depth images of 77 cows were used to train and test the model, and the dataset ratio was 1:1. The results showed that when the BCS value was in the range of 2.25 to 3.75 and the two-class classification limit was 2.75, the proportions of images with recognition errors less than 0.25 and 0.50 were 91.21% and 97.60%, respectively, and the average recognition speed was 3.446 frames/s. The recognition effect was better than that of the MobileNet-V2, XceptionNet and LeNet-5 network models. This method could realize the noncontact assessment of individual body conditions of dairy cows in large-scale farms and had the characteristics of high accuracy, strong applicability and low cost.

1. Introduction

Preparing a certain amount of fat before calving is an important guarantee for dairy cows to safely overcome the negative energy balance period, and evaluating body fat storage is a necessary prerequisite for ensuring the nutrition and health of dairy cows in the perinatal period (Martin and Ehle, 1986). The body condition score is a quantitative evaluation of animal fat retention and has become an effective tool to measure energy reserves, health status and feeding management level (Huang, 2020, Petrovska et al., 2014). The body condition score plays an important role in the prenatal and postnatal management of dairy cows, and abnormalities and mutations in body condition are also one of the

main characteristics of metabolic disorders in dairy cows (Rathbun et al., 2017, Chebel et al., 2018). Scoring is an effective way to prevent metabolic diseases, protect individual welfare and increase economic benefits to evaluate the body condition of dairy cows regularly.

At present, the body condition score of dairy cows mainly adopts artificial methods. According to certain rules, the fat accumulation degree of dairy cows is subjectively evaluated by visual observation of key areas such as the lumbar and sacrum (Roche et al., 2004, Zhao, 2017). However, due to the influence of artificial subjectivity, the repeatability and reliability of scoring results are low, and the evaluation process is time-consuming and laborious and relies heavily on the experience of evaluators. Therefore, the manual body condition scoring method has

* Corresponding authors at: College of Agricultural Equipment Engineering, Henan University of Science and Technology, Luoyang 471023, China (Jiangtao Ji). College of Electrical Engineering and Information, Northeast Agricultural University, Harbin 150030, China (Baisheng Dai).

E-mail addresses: jjt0907@163.com (J. Ji), bsdai@neau.edu.cn (B. Dai).

difficulty meeting the needs of real-time monitoring of the individual nutritional status of dairy cattle in large-scale farming (Kristensen et al., 2006, Rodríguez Alvarez et al., 2019, Song et al., 2019). An automatic system that can accurately score the body condition of dairy cows for a long time and continuously is urgently needed.

The application of machine vision in body condition scoring has gradually become a research hotspot, which is of great significance for reducing subjectivity and improving evaluation efficiency (Zhao, 2017, Huang et al., 2019). Bewley et al. verified the feasibility of applying machine vision to body condition scoring for the first time based on the contour and shape parameters of dairy cows calculated by 23 manual marking points in colour images and achieved a scoring accuracy of 92.79 % within 0.25 steps (Bewley et al., 2008). On this basis, Azzaro et al. developed an automatic algorithm to extract the above 23 feature points and used linear and polynomial kernel principal component analysis methods to reconstruct and test the shape descriptor of dairy cows. Validation testing showed that the average error of the polynomial model was 0.31 from the manual scores. (Azzaro et al., 2011). To further improve the degree of automation, Bercovich et al. extracted and analysed the curve of the cow's tail root contour and obtained a regression coefficient between the predicted BCS value and the real value of 0.77 under the condition of no interference (Bercovich et al., 2013). Wu et al. used the Hough transform to extract the tail region from the rear view of the cow's tail and removed the black and white flowers on the surface through an image block. Finally, the kernel-principal component analysis method was used to achieve a body condition recognition rate of 96.7 % within 0.5 errors (Wu et al., 2013). In the above scoring methods, although researchers initially realized the automatic evaluation of the body condition based on RGB images, the extracted geometric features of the body surface were mainly indirect features with weak correlation. In addition, due to the influence of the environment on image acquisition quality, the system had a low degree of automation, and it was difficult to meet the actual needs of aquaculture management in terms of real-time performance, recognition accuracy and reliability.

With the development of 3D machine vision technology, depth images have gradually become the mainstream technology of automatic scoring systems for cow body condition because they contain more features directly related to body condition and have great potential in reducing the subjectivity of scoring, improving the efficiency and accuracy of evaluation, and accelerating the commercialization of the system (Fischer et al., 2015, Imamura et al., 2017, Alborno et al., 2022). Spoliansky et al. developed an automatic body condition scoring system using a Kinect camera and established a regression prediction model using 14 eigenvalues, including breeding information (weight, age, etc.). The results showed that 91 % of the cow scoring error was within 0.5 steps (Spoliansky et al., 2016). Song et al. built a multi-view depth image system to extract the structural characteristics of eight body regions of dairy cows and used the nearest neighbour classifier to achieve a body condition score sensitivity of 0.72 within a 0.5 step size (Song et al., 2019). Zin et al. proposed two mathematical models in two regions of interest by extracting 3D surface roughness parameters. Based on the measurements of model accuracy, the automatic evaluation of BCS was realized, with a mean absolute percentage error (MAPE) of 3.9 % and a mean absolute error (MAE) of 0.13 (Zin et al., 2020). Zhao et al. constructed a variety of analysis models by positioning and analysing key bone areas, such as lumbar horn bone and hip bone, in deep images of the backs of dairy cows. Finally, 89 % of the sample score error was within 0.25 (Zhao et al., 2020). Combining the depth image with the supervised learning algorithm, the image was directly classified and recognized to realize the body condition score, which could omit the process of extracting the body condition eigenvalues. The accuracy rate of scoring within 0.25 errors can reach 82 % (Rodriguez et al., 2018). The use of deep learning and 3D technology further improves the accuracy of automatic body condition scoring. However, in actual production, to meet the nutritional management needs of dairy cows at

different growth stages, the gap between the body condition scoring value and the ideal value of dairy cows should always be maintained within 0.25 (Bewley et al., 2008, Zhao, 2017), while the accuracy of the existing automatic scoring system still has a certain disparity with the ideal standard of actual breeding management.

Therefore, to further improve the automatic scoring accuracy of cow body conditions, this paper took the depth image of the cow's back as the research object and proposed an automatic scoring method based on the EfficientNet network and convex hull features. First, the acquired depth image of the cow's back was preprocessed to extract the point cloud data of the back half body (lumbar horn bone to needle bone area) of the cow containing the main body condition information. Second, the convex hull feature image was obtained by voxelization and convex hull processing of the point cloud. Finally, a two-level model based on the EfficientNet network was constructed to classify the feature images and realize the accurate and efficient scoring of the individual body conditions of the cows in the breeding environment.

2. Materials and methods

2.1. System setup

The data for this study were collected at the University of Kentucky Coldstream Dairy Research Farm. A group of 77 Holstein cows was milked twice a day in the morning (5:00 AM) and afternoon (4:00 PM). After milking, all cows returned to the free-stall barn through the indoor narrow walkway. The width of the walkway was 1.03 m, which could effectively limit the activity amplitude and the number of single passes of cows and facilitate subsequent image acquisition and processing.

A PrimeSense™ Carmine 1.08 An RGB + depth sensor (Prime-Sense™, Tel Aviv, Israel) was used to capture the depth image of each cow's back when they walked freely through a narrow walkway. The camera system was installed on the ceiling 3.05 M above the concrete floor, and its field of view could completely cover the width of the walkway. The image was automatically captured by using the depth channel to monitor the foreground motion and stored on a computer hard disk through a 30 m USB2.0 extension cable connected to a computer in the dairy office. The resolution of the image data was 320 × 240 pixels, with a frame rate of 30 fps. The processor of the image processing platform was an Intel i5-8400, the main frequency was 2.80 GHz, the memory capacity was 8 GB, the operating system was Windows 10, and the algorithm development platform was MATLAB2020b.

An automatic video acquisition and control software was developed to detect the presence of cows in the scene and control the start and stop of image acquisition. As the cow walked from the left to the right of the field of vision, the background subtraction algorithm was used to obtain a binary image containing the cow. The four fixed lines in the image scene were used to trigger and stop the storage of the images; that is, when the cow's head (the rightmost pixel in the binary image) reached area between the third and the fourth line, the program started to store images; when the cow's tail (the leftmost pixel in the binary image) reached the area between the first and the second line, the storing of the images stopped. At the beginning of the recording, the software saved the initial background image (without a cow in it) for background subtraction later.

The purpose of setting four vertical lines is to reduce the storage of useless data and the impact of image distortion on data. When the cow's body appeared as much as possible in the field of view, it began to store images, which greatly reduced the amount of data for subsequent processing. As shown in Fig. 1, the four vertical lines are four fixed columns in images, and the distances of the first and second line from the left side of the image are approximately 10 % and 25 % of the image width, respectively. Lines 3 and 4 are symmetrical to lines 2 and 1, respectively, and are located in the right half region of the image.

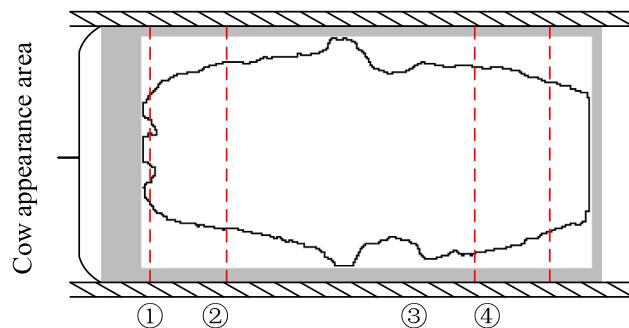


Fig. 1. Positions of the four detection lines during image acquisition.

2.2. Data acquisition

Depth images of 77 lactating cows were collected twice a day after milking from April 1st to June 7th, 2014. Over this time, three experienced professional scorers manually scored all of the experimental cows once a week for the same period of time according to the same scoring standard (5-point system). The median of the three scores from one cow was used as the body condition score of the cow in the corresponding week to reduce the subjective difference of the manual score. Studies have shown that the change in the body condition score of dairy cows in a week under normal conditions is not more than 0.25 (Zhao, 2017). Therefore, the scores of each cow over consecutive weeks were compared to remove outliers. The built dataset contained 5119 depth images of 77 cows and the corresponding manual scoring values. A total of 2627 images of 39 cows were randomly selected for model training, and 2504 images of the remaining 38 cows were used for model testing, with no crossover between the training set and the test set data. Table 1 shows the BCS range and sample distribution of the two datasets. For the extreme categories with fewer samples (≥ 4.00 and ≤ 2.25), Ferguson et al.'s opinion was adopted to classify them into 3.75 and 2.25 categories, respectively.

2.3. Image processing

To facilitate the subsequent algorithm development and feature extraction, the original depth image is preprocessed by target extraction, target rotation and hindquarters image acquisition to obtain the point cloud on the back of the cow hindquarters with the main body condition information. The specific steps are shown in Fig. 2.

2.3.1. Target extraction

To improve the quality of target extraction, a batch of depth frames without cows were used to build initial background image. The background image was then continuously updated to avoid potential error from any single background image. The depth image of cows without background could be obtained by a difference operation between the current image and the background image obtained by modelling, and threshold processing was carried out on the difference result. After background subtraction, the depth values of the cow image pixels were

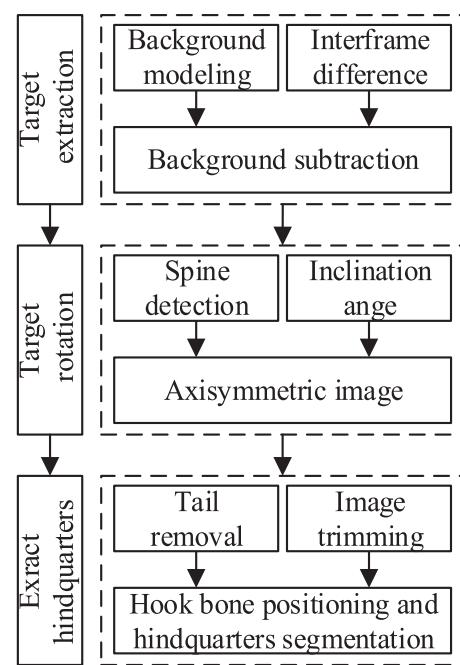


Fig. 2. Image preprocessing steps.

converted to the distance from the floor by adding the camera height (3050 mm).

The detailed process for background modelling is shown in the Fig. 3. When the camera starts to collect images for the first time, 1200 background images without cows are collected before the cow enters the walkway, and these images are averaged to obtain background image for subsequent target detection. For the new image captured by the camera, firstly, the difference image between the current frame and background is calculated, and processed with a threshold to get the area with large difference. If the areas with large difference are not found, the image is regarded as a new background and added to the background image sequence. Meanwhile, the earliest image in the original sequence is deleted and averaged background image is updated. If an area with large difference is found, it is necessary to further judge whether the area is within the range where cows may appear. If the judgment is yes, it indicates that the cow target appears in the image, and enter the following processing to detect cow target. If the area is not in the range where the cow may appear, it indicates that an abnormal target intrusion occurs in the image which will be discarded. In this study, the camera was installed above the cow's walkway, and there were fences on both sides of the walkway to limit the cow's walking area. The range where cows might appear is shown in Fig. 1 in the manuscript.

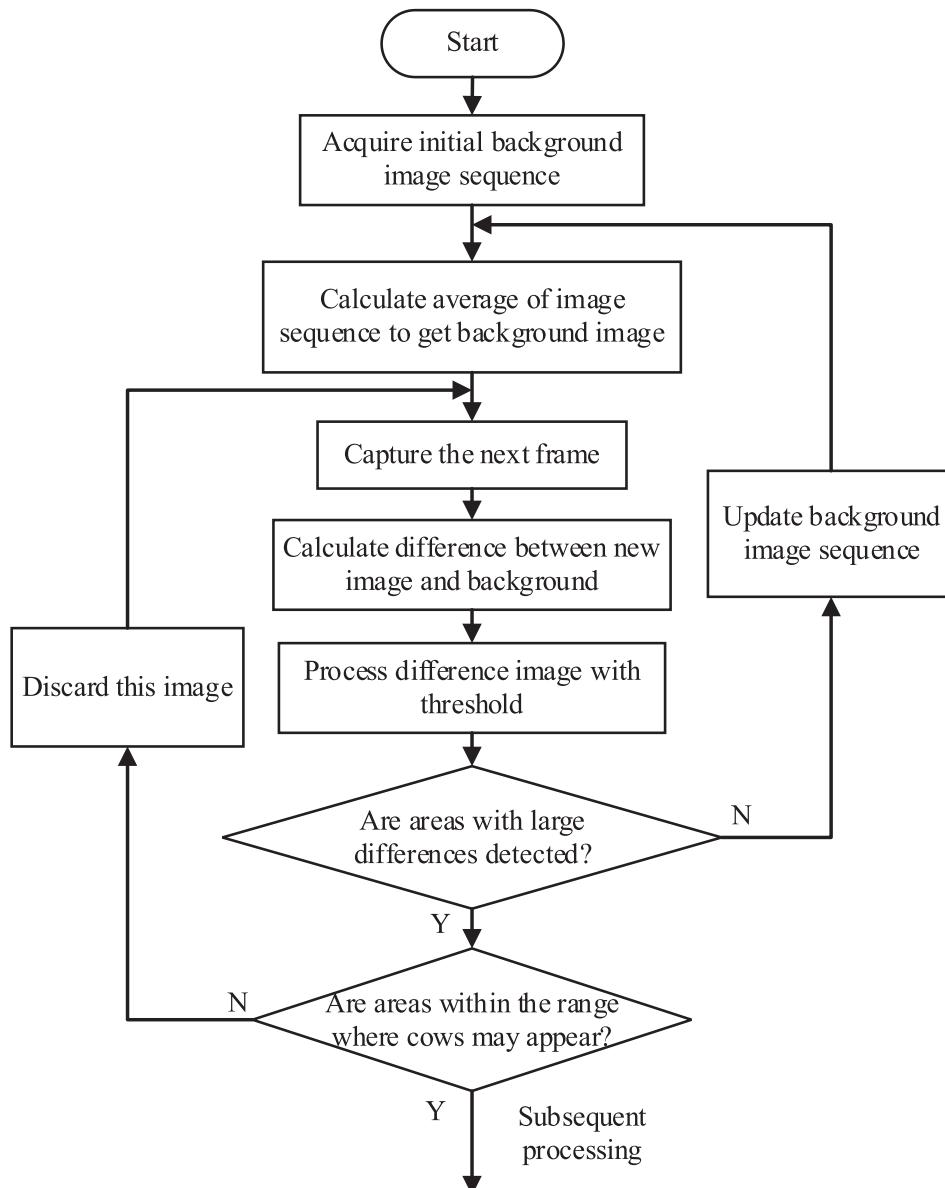
2.3.2. Target rotation

To reduce the influence of cow body tilt on data consistency in the acquisition process, the extracted target cow image was rotated to make its spine direction parallel to the horizontal axis of the image so that the subsequent algorithm could perform symmetry processing on both sides along the spine direction. In this study, the images were taken during the cows returned to the free-stall barn through the narrow walkway after milking. Cows walked in one direction when passing through the walkway, as they appeared from the left side and walked right through the field of view. Therefore, the direction of the cow's spine is fixed, that is, the left side is the tail and the right side is the head. Of course, we can also judge the walking direction of the cow by analyzing the consecutive frames, and then recognize the direction of the cow's spine.

The maximum value of each column in the depth image was selected as the spine point, and the spine line was obtained by the linear fitting method. The image was rotated according to the angle between the

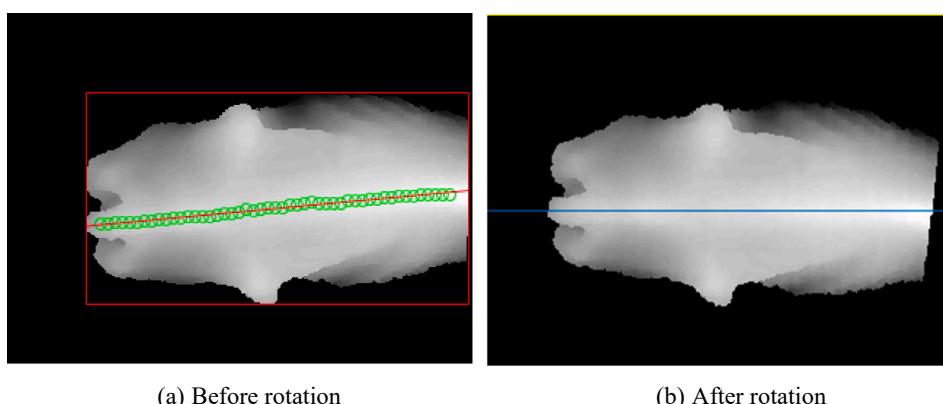
Table 1
BCS value distribution over the training validation and test sets.

BCS value	Distribution of training images	Distribution of test images
2.25	243	214
2.50	988	849
2.75	792	872
3.00	212	279
3.25	158	82
3.50	125	126
3.75	109	82

**Fig. 3.** Process of background modelling and updating.

fitted line and the transverse axis of the image. The symmetry of the rotated image was tested by calculating the second moment of the cow target. During walking, the postures of cow's head and neck are more

flexible and prone to twisting the spine. The distance between the left-most pixel from the cow target and the left side of the image was determined. When the distance was greater than a value, the

**Fig. 4.** Fitting the spine and image rotation.

hindquarter area of the cow fully entered the field of view. At this time, due to the width of the image field, the cow's head and neck were generally outside the image. To further avoid the effect of the head and neck on line fitting of spine, the cow target was analysed to segment the cow head and neck area according to the width of the pixels in each column, and only the cow's back area was remained. Fig. 4 shows the detection of cow spine points and the result of image rotation. There was still a certain torsional deformation between the forequarters and the hindquarters of the body divided by the sacral ligament. But the torsional deformation between the forequarters and the hindquarters of the cows during walking is relatively small, and the effect on the linear fitting is not obvious.

2.3.3. Extract hindquarters

Since the investigation areas of body condition score (hook bone, needle bone, spine, short ribs, hip bone and tail bone) are all located in the hindquarters of dairy cows, the forequarters image without body condition information was deleted to reduce the redundancy of image information, and the discrete pixel points and tail (influence score) in the hindquarters image were removed to improve the image quality. The results are shown in Fig. 5, and the detailed steps were as follows:

1) **Tail removal and filtering.** The column in which the pixels on the leftmost side of the image were located was identified, and starting with this column, when the span $W_j(j = 1, 2, \dots, n)$ of all non-0 pixels in a column was greater than the threshold T_t (set in the experiment to 200), the pixel value of this column was revised to 0 so that the image was split into two parts (Fig. 5a). Then, according to the connected domain properties of the image region, the tail and discrete pixel points in the posterior abdomen were filtered out (Fig. 5b-c).

2) **Image pruning.** A depth threshold D_T was used to segment the region of interest associated with the body condition score in depth images (Fig. 5d). Measuring from the back to the front of the cow, the pixels for which the difference between the pixels in each column and the maximum pixel value of the column was greater than a certain threshold were removed, and the remaining pixels were retained. After repeated experiments and verification, it was found that the image trimming effect was the best when $D_T = 100$ mm.

3) **Hook bone localization and hindquarters segmentation.** The contour of the back of the cow was divided into left and right parts by a symmetry line. As shown in Fig. 5e, points A and B, farthest away from the symmetry line, indicate the ends of the hook bones on the left and right sides of the cow's outline, respectively. The x-coordinates of points A and B indicate the x-coordinates of the hook bones.

The hook bones are the bumps on the left and right sides in the curve of the sacral ligament. The sacral ligaments were isolated by connecting points A and B and extracting the slice on the line from the depth image. Because the sacrum ligament was the obvious demarcation line between the forequarters and hindquarters of the trimmed cow, the hindquarters of the cow (Fig. 5f) could be obtained by saving the pixels on the rear side of the column where the demarcation line was located. To facilitate subsequent feature extraction, the cow hindquarters image was converted to a point cloud.

2.4. Convex hull feature extraction

The body condition score of dairy cows mainly consisted of a comprehensive evaluation of subcutaneous fat thickness in each investigation area, and the fat filling degrees of dairy cows with different fat and lean degrees were different in different investigation areas. Therefore, the hollow space formed by the body surface with different fat filling degrees and the surrounding bone convex hull were also different (Fig. 6). For fat dairy cows, due to the accumulation of fat, the depression in each region was small, and the hindquarters were centred on the spine, with a flat transition to both sides. Because of its low fat content, bone (spine) is more prominent, the overall depression is obvious, and

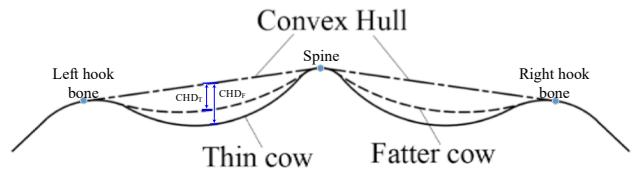


Fig. 6. Convex hull distance of the sacral ligament of thin cow vs fat cow.

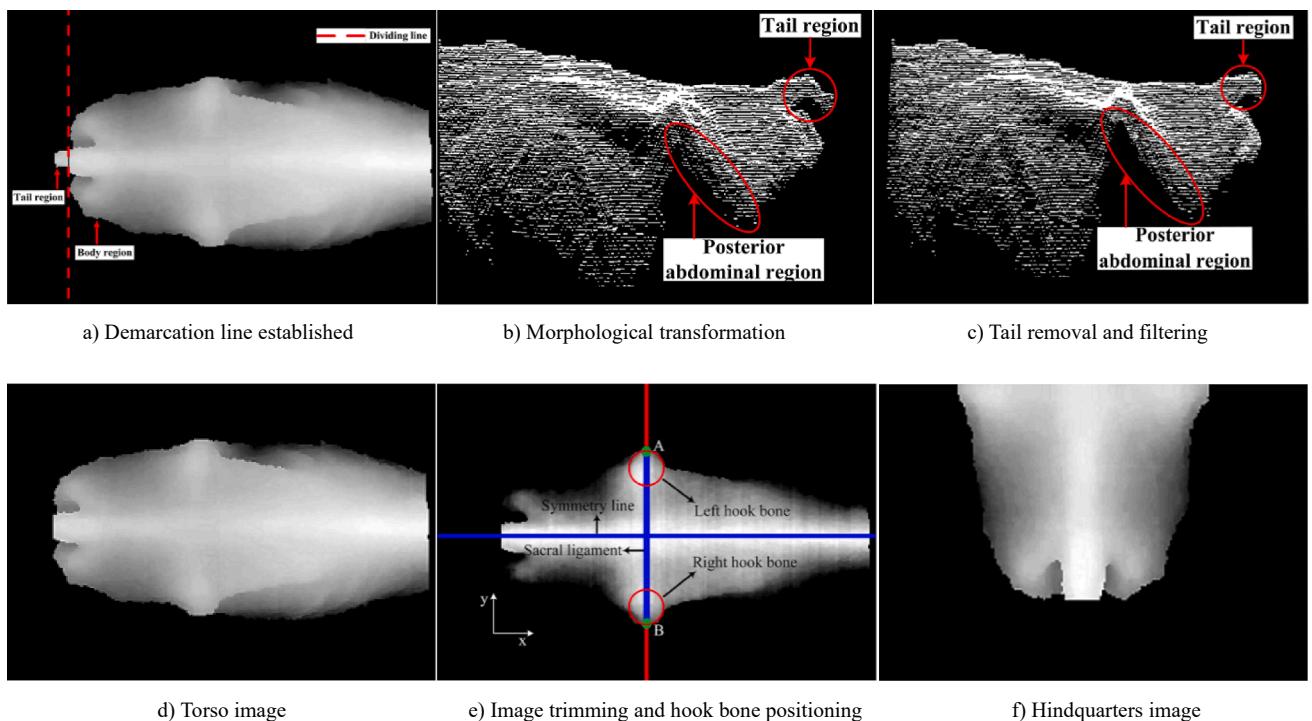


Fig. 5. Process for extracting image of cow's hindquarters.

the back shows an inverted V type.

Convex hull distance is the distance between any point and the convex hull in the vertical direction. As shown in Fig. 6, the curve of sacral ligament was aligned based on the hook bones and spine of the cow, so that the difference of convex hull distances between the thin and fat cows is significant. In the hindquarter region of the cow, the alignment of key points makes the distribution of convex hull distance comparable.

The process of a cow shifting from lean to fat involves the accumulation and filling process of body fat in the region around bones. Therefore, based on the convex hull principle and the gradual change characteristics of the cow's body shape from lean to fat, a three-dimensional feature to quantify the sagging degree between bones within cow's hindquarter area was proposed based on the convex hull analysis. The specific steps were as follows:

1) **Statistical filtering.** As shown in Fig. 7, the statistical filtering method was used to filter out the sparse outlier point cloud on both ventral sides of the hindquarters of cows to obtain the compact point cloud subject to Gaussian distribution. The mean value of the Gaussian distribution was determined by the average distance between a point and its 200 adjacent points, and the standard deviation coefficient was set as 3.

2) **Voxelization.** Affected by the individual size difference of cows and the body shaking of cows during the shooting process, the point clouds of hindquarters of cows obtained were not comparable in terms of order of magnitude and the information content of each coordinate data value. Therefore, inspired by the pixel position serial number and quantitative characteristics of the image, voxel gridding was carried out on the filtered hindquarters body point cloud (Fig. 8) so that it initially had the image characteristics. The length, width and height (L, W, H) of the initial voxel were determined by the ranges of the point cloud coordinates in XYZ axis, and the voxel resolution was set to (10, 10, 10). The coordinates of voxel barycentre were the mean of the coordinates from all points located within the voxel. After voxel gridding, the voxel numbers were (i, j), and each voxel contained only one point. If the x, y coordinate value of each voxel barycentre was replaced by the voxel numbers and its z coordinate value was taken as pixel values, the voxel cloud would initially have image characteristics and be comparable in the x, y, z directions through subsequent size transformation and pixel value replacement.

3) **Feature extraction of convex hull.** Based on the influence of fat accumulation degree on the shape of cow hindquarters, the three-dimensional convex hull of the point cloud after voxelization was calculated, and the distance between the peripheral voxel and the nearest convex hull surface was calculated and projected onto the X-Y

plane to obtain the convex hull feature map to identify the fat and thin degrees according to the difference of the sagging degree of different cow's complete hindquarters. The hollow of the whole hindquarters was represented by the feature vector composed of the nearest distance (feature distance) between each point inside the convex hull polyhedron (including the boundary point) to its surface. The feature distances of invalid voxels were set as 0. The X, Y and Z coordinate values of each voxel centre of gravity were replaced by its voxel number i, j and feature distance, and the feature image was formed. The replacement of coordinate values conferred comparability and body condition evaluation significance in the Z-axis direction. However, affected by the difference in the length and width of the initial voxel, when the voxel resolution was the same, the sizes of the feature images formed by different hindquarters point clouds were not all the same. Therefore, the nearest neighbour interpolation method was used to transform the size so that the feature images of different point clouds were comparable. The size of the feature images in the test process was unified to 60 pixels × 60 pixels. The calculation of the convex hull used the convhulln function in MATLAB, and its core idea was based on the Quickhull algorithm (Barber et al., 1996). The Quickhull algorithm is a geometrical method. Firstly, the Quickhull algorithm initializes a tetrahedron composed of four points, then divides the point set into the external point set of each surface, and selects the point furthest from the surface for processing (Barber et al., 1996). According to the Beneath-Beyond theory, a face either remains unchanged or creates new faces based on the theory. The algorithm can generate the convex hull of the processed points, and continuously add new points for processing until the points on the point set are processed. This algorithm did not need to set parameter values and had good robustness and efficiency. The 3D convex hull was composed of a series of known vertices in the 3D point set Q (the number of points was greater than 4), so in the process of forming the convex hull, it was not affected by the characteristics of the data itself (whether it was random data, whether there was noise). The characteristics of the data only affect the computation time of the algorithm, and compared with the gift wrapping, Clarkson and Shor algorithms, the Quickhull algorithm was more efficient. In this paper, the filtering before the convex hull calculation was performed to filter out the noise in the hindquarters point cloud to avoid the potential noise points causing one side of the convex hull to be extremely prominent and thus affecting the final quality of the convex hull. Fig. 9 shows the convex hull calculation results and convex hull feature diagram. Through manual inspection, the convex hull and feature map of all point cloud data are correctly calculated.

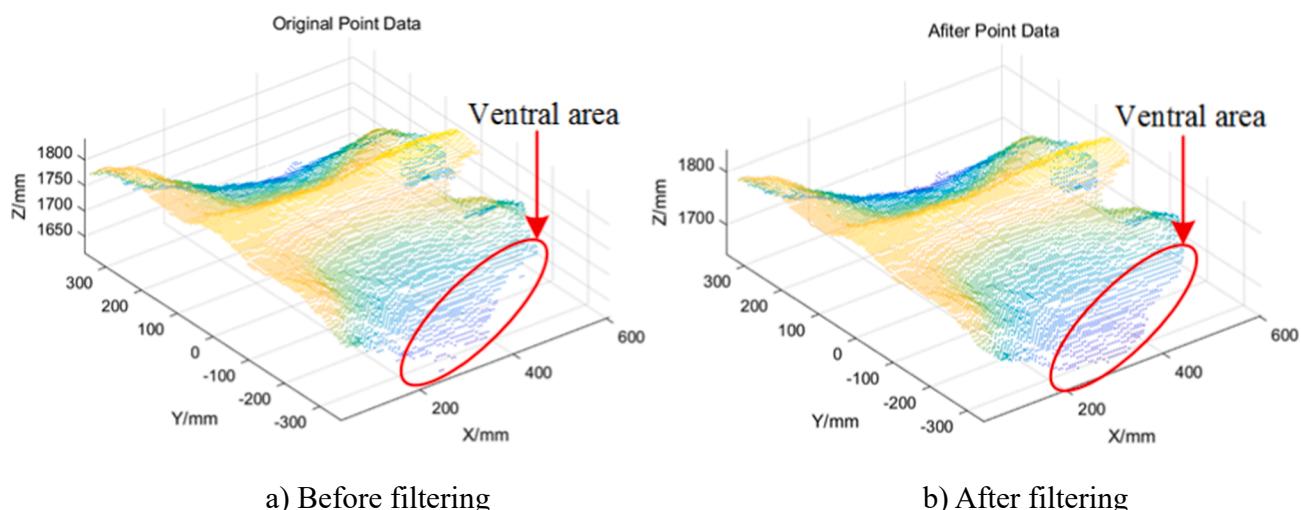


Fig. 7. Statistical filtering of hindquarters point cloud.

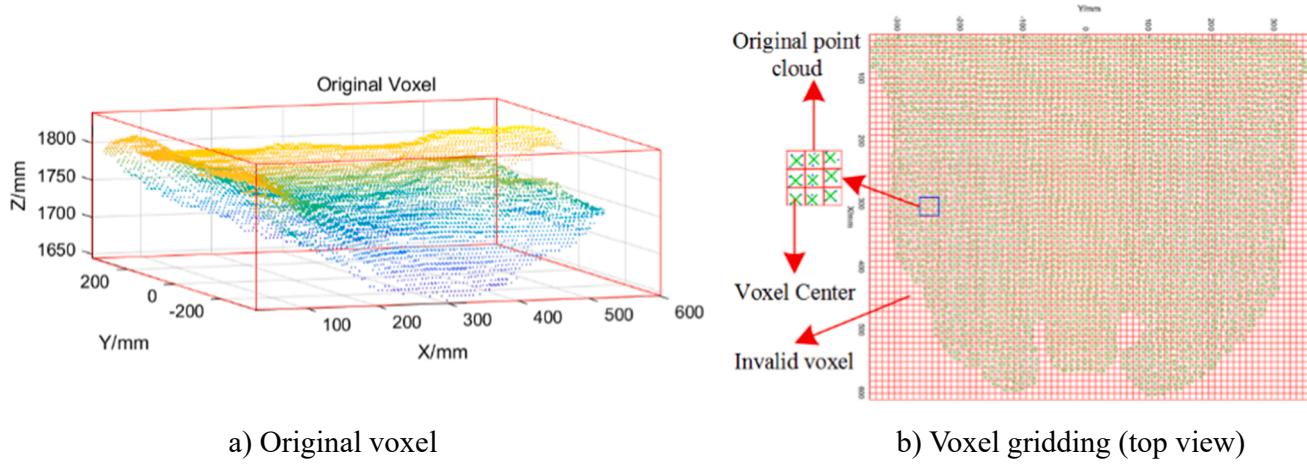


Fig. 8. Voxel gridding of hindquarters point cloud.

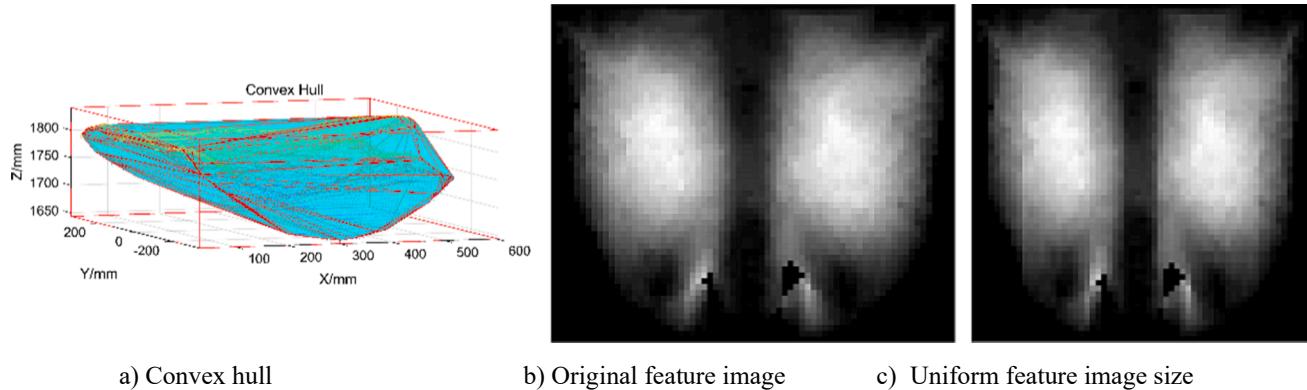


Fig. 9. Convex hull feature extraction and visualization.

2.5. Prediction models

2.5.1. Secondary model

Based on convolutional neural networks with obvious advantages and wide applications, this paper constructed a two-level model for feature image classification to simplify the cumbersome process of feature design and extraction in traditional image classification methods, as well as the generated suboptimal accuracy and efficiency, and improve the automation level and accuracy of the BCS scoring process. The two-level model provides a powerful tool to solve the problem of low classification accuracy of low-sample categories caused by the uneven distribution of samples among BCS classes, improves the size and efficiency of the classification model to be suitable for mobile or

embedded devices, and avoids the possible underfitting of low-sample categories in the direct prediction model. The structure is shown in Fig. 10. First, 2627 feature images of the training set were classified by the M1 network with a certain category X as the boundary. Then, M2 and M3 networks were used to finely classify the two types of results after classification by BCS (2.25 – 3.75). Finally, the accuracy of the BCS secondary prediction model composed of M1, M2 and M3 networks was tested with 2504 test set feature images as input. M1, M2 and M3 were the same network EfficientNet. In addition, to avoid the influence of category boundaries on the final model test accuracy, this paper compared and analysed the two-level model accuracy under three different category boundaries.

2.5.2. EfficientNet

The EfficientNet network adopted a simple and efficient compound scaling method. Under the constraints of preset memory and computation size, a set of fixed scaling coefficients (α , β and γ) was found by grid search to uniformly scale the baseline network dimension (depth d , width w , resolution r) to achieve the balance of network dimension and the improvement of overall performance of the model (Tan and Le, 2019). In addition, the network developed the baseline network (Efficientnet-B0) by leveraging a multiobjective neural architecture search that optimized both the model performance and the effectiveness of its expansion. As shown in Fig. 11(a), the network structure of Efficientnet-B0 consisted of seven parts, with each comprising a “block.” First, feature images were passed through a 3×3 convolution layer to generate partial features of feature images. Rich feature details for prediction were then further extracted in several block structures. Finally, the convolution layer-pooling-full connection operation was

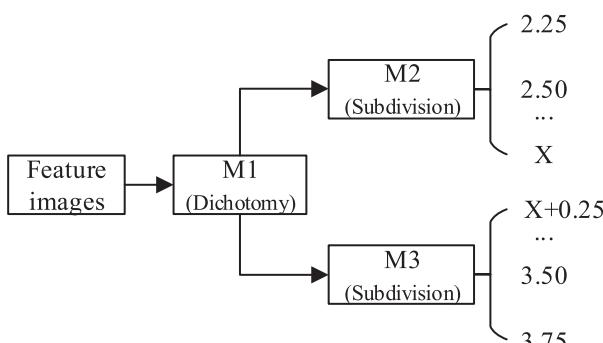


Fig. 10. Structure diagram of two-level model.

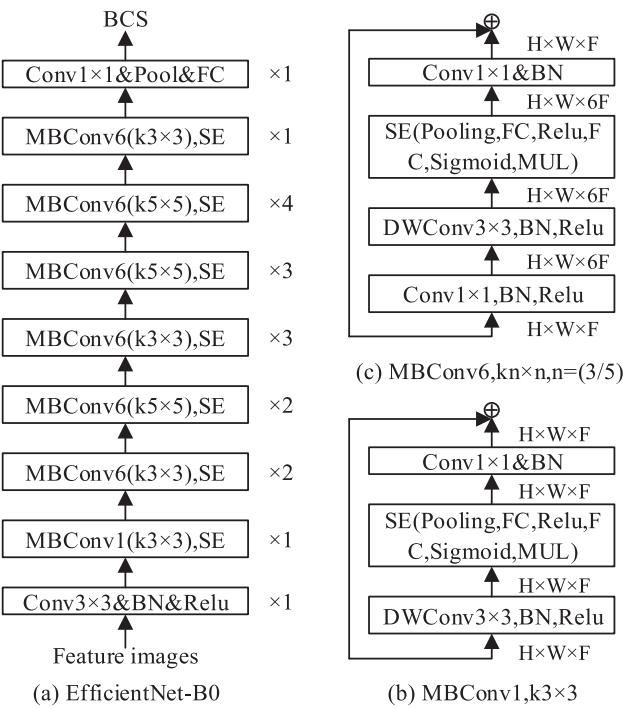


Fig. 11. The structure diagram of the Efficientnet-b0 network.

applied in place of a classifier to differentiate the BCS category represented by the feature image. Among these operations, three block structures were combined by connecting the fronts and ends of a few identical mobile inverted bottleneck conv (MBConv) outputs. Squeeze-and-excitation (SE) was added for optimization so that the shallow neural network could capture image features via a global receptive field and provide descriptions for the images. In addition, to enhance the nonlinearity and network expression ability of features in high-dimensional space and avoid gradient disappearance during model training, the ReLU function is selected as the activation function of the network, and the softmax regression function and cross entropy loss function are used to normalize the output of the fully connected layer and calculate the loss. H, W and F represent the height, width and number of channels of the input image, respectively.

The precision of the model can be maximized by balancing network dimensions under the condition of limited computational ability. To capture fine-grained images and improve the generalization effect on body condition classification tasks, a compound-scaling method was selected for use in the benchmark network. To make the network model more suitable for convex hull feature datasets and achieve optimal classification performance, the whale optimization algorithm (WOA), which is widely used in multi-objective optimization, was used to obtain the optimal values of α , β and γ during the grid search of the baseline network. The algorithm was simple to operate, had few adjustment parameters, had a strong ability to jump out of local optima, and had a high calculation accuracy (Mirjalili et al., 2016). The process was as follows: first, a group of uniformly distributed random solutions $G(0) = \{\bar{X}_1(0), \bar{X}_2(0), \bar{X}_3(0), \dots, \bar{X}_i(0), \dots, \bar{X}_N(0)\}$ was created to form the initial population, where $\bar{X}_i(0) = (\alpha_i, \beta_i, \gamma_i)$, $i = 1, 2, \dots, N$, and initialized the population number N , the maximum number of iterations T_{max} , and the value range of α , β and γ . The test accuracy of the baseline network was set as the fitness function $F(x)$. Second, the individual fitness value was calculated, and the position where the optimal fitness value was located was taken as the optimal individual position x_{best} . Then, when the coefficient vector $|A|$ greater than 1, the next generation position was updated according to the random search mechanism. When $|A| \leq 1$, guided by the optimal individual position in the current population, the

next generation position was updated by selecting the shrinking encircling mechanism or the spiralling updating mechanism with 50 % probability. The steps were repeated until the maximum number of iterations was met. According to the above steps and the initialization parameters shown in Table 2, the scaling coefficients of the baseline network were optimized on 5119 convex hull feature images. As shown in Fig. 12, the optimal scaling coefficients of α , β and γ were 1.45, 1.35 and 1.1, respectively. To further improve the performance of the network model after composite scaling, $\varphi = 7$ was set in the model scaling process to obtain the EfficientNet-B7 network model. The initialization parameters of network model training were set as 0.001 initial learning rate, 100 maximum iterations, 0.9 momentum, 30 verification frequency and 64 minimum batch size.

2.5.3. Performance evaluation

As a powerful tool to analyse the ability of a given classifier to recognize different class tuples, a confusion matrix has been widely used in the evaluation of classification methods because it provides more information than a single evaluation index (accuracy or precision). Therefore, this paper uses the confusion matrix to manifest the classification details of the model for each BCS category, and based on this, the following metrics were calculated:

Accuracy: effectiveness of a classifier, that is, the percentage of samples correctly classified.

$$\text{Accuracy} = \frac{\text{correctPredictions}}{\text{predictions}} \quad (2)$$

Precision: ability of the classifier not to label a negative example as positive. Since the body condition was a multiclassification problem, it was defined as the mean precision of all classes.

$$\text{Precision} = \frac{\sum_{i=1}^{\text{classes}} (tp_i / (tp_i + fp_i))}{\text{classes}} \quad (3)$$

where tp_i and fp_i are the number of correct recognition and error recognition in prediction class i , respectively, and classes are equal to the number of categories, 7.

Recall: ability of the classifier to find all the positive samples. It was defined as the mean recall of all classes.

$$\text{Recall} = \frac{\sum_{i=1}^{\text{classes}} (tp_i / (tp_i + fn_i))}{\text{classes}} \quad (4)$$

where tp_i and fn_i are the number of correct recognitions and error recognitions in true class i , respectively.

F1-score: The harmonic mean of precision and recall rate, which was used to measure the benignity of the classifier in the presence of rare classes.

$$\text{F1-score} = \frac{\sum_i^{\text{classes}} \left(2 \times \frac{\text{Precision}_i \times \text{Recall}_i}{\text{Precision}_i + \text{Recall}_i} \right)}{\text{classes}} \quad (5)$$

Table 2
Initialization setting of WOA optimization algorithm parameters.

Control Parameter	Values
Dimension D	3
Whales Population N	200
Maximum Iteration T_{max}	50
a^1	$2 \rightarrow 0$
b^2	1
The range of α , β , γ	[1,2]
Coefficient vector A	[-2,2]

1 a is a vector that linearly decreases from 2 to 0.

2 b is a constant for defining the shape of the logarithmic spiral.

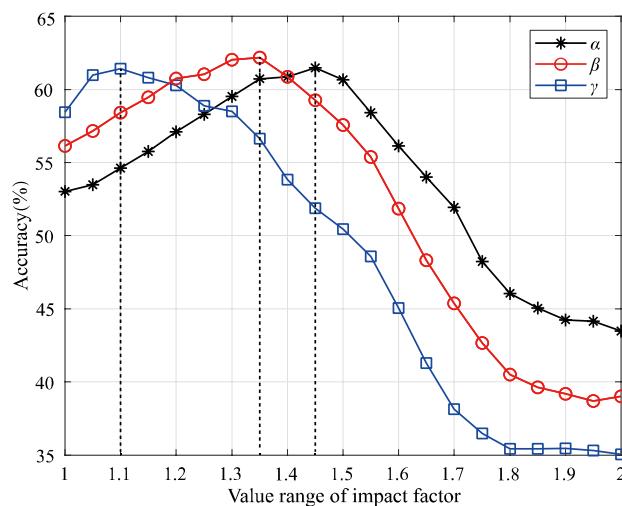


Fig. 12. Optimization results of scaling coefficients α , β and γ .

where $Precision_i$ and $Recall_i$ are the precision and recall of prediction class i and true class i , respectively.

3. Results

3.1. Performance of direct model

Before model construction, this paper first explored the scoring performance of the direct model and, inspired by it, proposed a second-level scoring model. Fig. 13 shows the confusion matrix of the scoring results of the direct model (EfficientNet). The three colours of red, orange and yellow represent accurate predictions, with predictions within 0.25 and 0.5 errors, respectively. Although it was different from the main diagonal distribution of the ideal confusion matrix, within the range of 0.50 errors, the BCS prediction results based on convex hull feature images showed an approximate ideal trend along the diagonal distribution. This indicated that cows with the same BCS value did not necessarily have the same prediction results, and the recognition accuracy of the model in the error range of 0, 0.25 and 0.50 was 45.88 %, 88.74 % and 96.33 %, respectively. In addition, compared with ≥ 3.00 class, the prediction results of the 2.50 and 2.75 classes are more

dispersed, but within the 0.25 errors, both of them achieved a greater than 90 % recognition rate, which is far better than the other categories. The low recognition rate of categories 3.00–3.75 might be related to the uneven distribution of samples between classes, and the model learned more high sample category features. Inspired by the difference in sample distribution and recognition rate, this paper trained classification models of high and low sample categories to improve the low classification accuracy of low sample categories to further improve the overall accuracy and explored the influence of different classification boundaries on the model accuracy.

3.2. Confusion matrix

Fig. 14 shows the confusion matrix of the scoring results of the second-level model when different BCS categories (2.75, 3.00, 3.25) are the classification boundaries. It was possible to see that more than 73 % (16 % higher than that of Rodriguez et al.) of samples per class in the above three cases were distributed to the diagonal and its two sides, except for category 3.25, which indicated that the proposed pattern recognition method based on convex hull features could effectively distinguish cow variability related to BCS changes. In particular, the recognition superiority (97 %, 90 %) of common BCS (2.50, 2.75) within the 0.25 errors range provided the possibility to greatly improve the body condition detection accuracy of most dairy cows in large-scale farming. The low-sample distribution problem of the main diagonal presented in the 3.25 class classification might be associated with the fact that the class was located at the fat and thin boundary because it was difficult for the model to extract relevant information associated with this value and form effective criteria that are different from those of adjacent categories considering only a few training samples. With the increase of the classification boundary from 2.75 to 3.25, the accuracy of dichotomies gradually increased from 87.5 % to 94.49 %, but the overall accurate recognition rate of BCS decreased from 45.01 % to 41.97 %. When the error range was expanded to 0.25 and 0.50 steps, the fluctuation trend decreased first and then increased, and the highest recognition accuracy was achieved when 2.75 was the boundary, which was 91.21 % and 97.60 %, respectively. In addition, compared with other boundaries, when 2.75 was the boundary, each class had a more obvious advantage in recognition accuracy within 0.25 errors. From the above analysis, it could be seen that when category 2.75 was used as the classification boundary, the classification performance of the second-level model was optimal. The recognition accuracy within 0.25 errors was 91.21 %, which was 2.74 % higher than that of the direct model, and the recognition rates of various categories were significantly improved, especially the 2.25, 3.00, 3.25 and 3.75 categories, which were increased by 7.01, 7.53, 5.38 and 13.42 percentage points, respectively.

3.3. Model evaluation

Table 3 shows the classification performance of the model in different error ranges with category 2.75 as the classification boundary. The precision of BCS less than or equal to 2.75 was significantly higher than that of BCS greater than 2.75 at 0 errors, which verified the conclusion of Sun et al. This trend was not affected by the parameter category at 0.25 and 0.50 errors. As the error range increased, the difference in F1 score (the harmonic mean of precision and recall) between thin ($BCS \leq 2.75$) and fat ($BCS > 2.75$) cows gradually narrowed, and the average differences were 0.22, 0.19 and 0.06 within the error ranges of 0, 0.25 and 0.50, respectively. In addition, the precision value within the error range of 0.50 did not significantly improve because the number of samples classified as $BCS = 3.75$ was small, which made the F1 score value within this error limited.

3.4. Model comparison

To verify the influence of different pattern recognition networks on

True Class	Predicted Class						
	2.25	2.50	2.75	3.00	3.25	3.50	3.75
2.25	92	95	26	0	1	0	0
2.50	44	473	317	7	6	1	1
2.75	10	305	410	27	56	13	6
3.00	0	27	135	31	50	23	13
3.25	0	4	41	9	23	3	2
3.50	0	1	8	13	16	70	18
3.75	0	1	0	7	15	9	50

Fig. 13. Confusion matrix of scoring results by EfficientNet network.

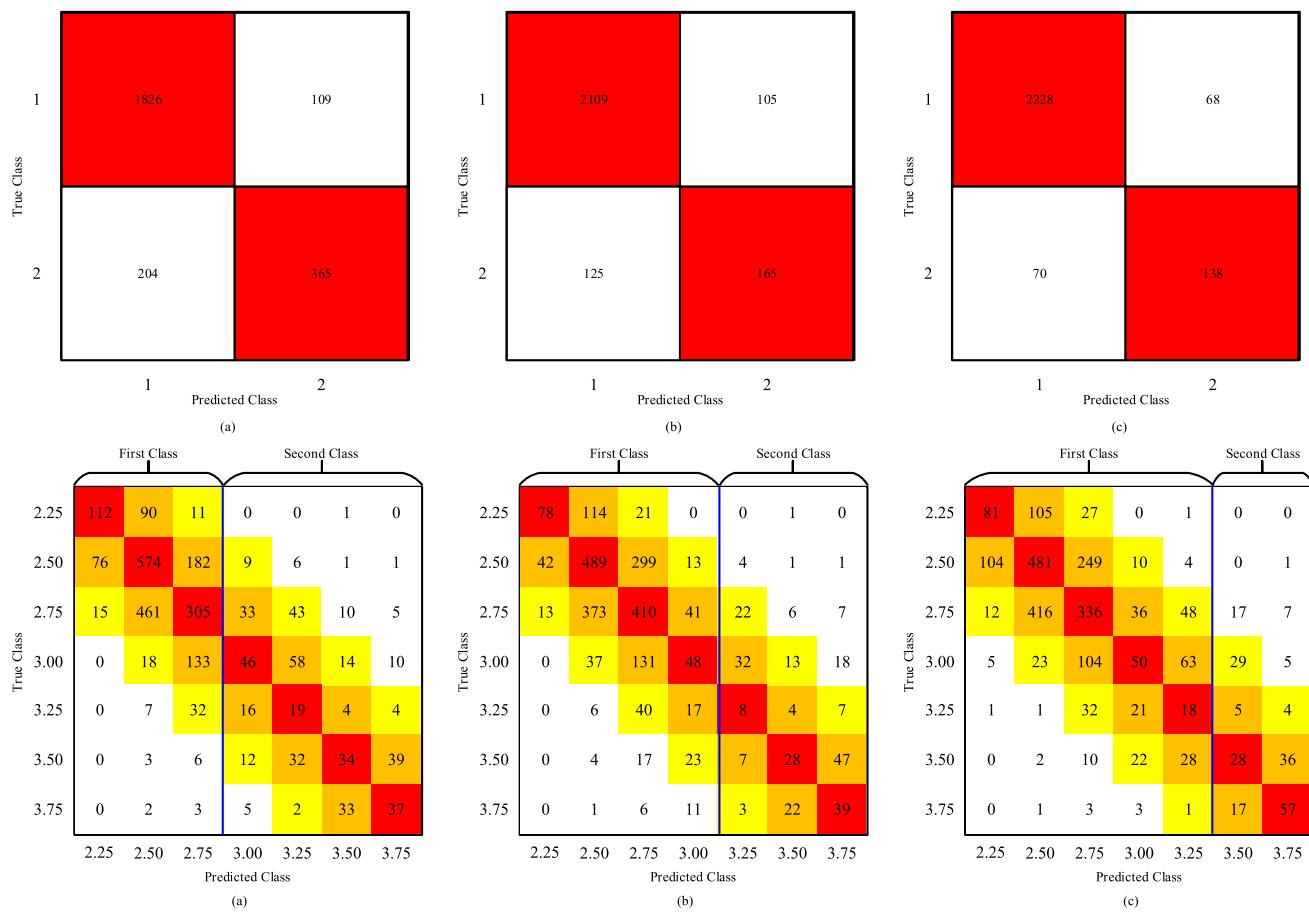


Fig. 14. Confusion matrix of the scoring results of the secondary predictive model under different boundaries.

Table 3
Classification measurement within different range errors.

BCS	0(exact)			0.25			0.50		
	Precision	Recall	F1-score	Precision	Recall	F1-score	Precision	Recall	F1-score
2.25	0.55	0.52	0.53	0.93	0.94	0.93	1	0.99	0.99
2.50	0.50	0.68	0.58	0.97	0.98	0.97	0.99	0.99	0.99
2.75	0.45	0.35	0.39	0.92	0.92	0.92	0.99	0.98	0.98
3.00	0.38	0.16	0.23	0.79	0.85	0.82	0.96	0.96	0.96
3.25	0.12	0.23	0.16	0.68	0.48	0.56	0.96	0.91	0.93
3.50	0.35	0.27	0.30	0.73	0.83	0.78	0.88	0.93	0.90
3.75	0.39	0.45	0.42	0.79	0.85	0.82	0.83	0.88	0.85
Average	0.39	0.38	0.37	0.83	0.84	0.83	0.94	0.95	0.94

the final score results when category 2.75 was taken as the classification boundary, the classification performance of four network models (LeNet-5, XceptionNet, MobileNet-V2, EfficientNet) was compared and analysed based on the convex hull feature image dataset. The confusion matrix is shown in Fig. 15. Overall, the confusion matrix of the four models showed little difference, and the error distribution between the predicted value and the real value was concentrated within 0.50 for the whole sample. However, for each category, the four models showed different characteristics. When the error step was 0.5, the four models all achieved the lowest and highest recognition rates in the 3.25 and 2.50 categories, respectively. In the categories of 3.00 and 2.50, the recognition rates of the four networks from low to high were LeNet-5, MobileNet-V2, XceptionNet and EfficientNet. However, the ranking changed over the 2.75 class, with the XceptionNet network performing the worst and significantly lower than the optimal EfficientNet. In the 3.50 category, the EfficientNet and LeNet-5 networks were superior to the MobileNet-V2 and XceptionNet networks. In addition, the order of

network recognition rates for BCS classes 2.25 and 3.75 was the opposite, and the recognition rates of the other three networks were all above 91 % in the 2.25 category except XceptionNet. The structural difference of the model leads to different learning abilities of the feature image, which was the main reason for the difference in recognition accuracy among BCS classes.

Due to the potential subjective errors of manual scoring and the nutritional management requirements of dairy cows (the gap between the predicted value and the ideal value of BCS was always maintained within 0.25 errors), it was difficult to give accurate breeding management decisions by the model parameters within 0 and 0.50 errors. Therefore, to further accurately analyse the model performance of the four networks and improve the guidance of the optimization model for farm feeding management, according to the confusion matrix of the sample classification of the test set, the indicators defined in Section 1.6 and recognition speed were adopted to comprehensively evaluate the four networks within the error range of 0.25. The evaluation results are

True Class	2.25	2.50	2.75	3.00	3.25	3.50	3.75	
Predicted Class	2.25	113	81	19	0	0	1	0
2.25	140	460	222	13	11	0	3	
2.50	31	400	307	76	33	7	18	
2.75	1	24	82	70	42	13	47	
3.00	0	1	29	16	17	1	18	
3.25	0	2	3	10	16	44	51	
3.50	0	1	10	0	0	12	59	
3.75	2.25	2.50	2.75	3.00	3.25	3.50	3.75	

(a) LeNet-5

True Class	2.25	2.50	2.75	3.00	3.25	3.50	3.75	
Predicted Class	2.25	87	97	29	0	0	1	0
2.25	68	510	251	8	9	2	1	
2.50	19	347	384	37	63	9	13	
2.75	1	19	100	74	50	13	22	
3.00	0	0	31	22	18	2	9	
3.25	0	1	2	27	24	26	46	
3.50	1	2	0	0	1	16	62	
3.75	2.25	2.50	2.75	3.00	3.25	3.50	3.75	

(b) XceptionNet

True Class	2.25	2.50	2.75	3.00	3.25	3.50	3.75	
Predicted Class	2.25	27	180	6	0	0	1	0
2.25	5	593	227	9	11	4	0	
2.50	1	339	382	53	75	18	4	
2.75	0	21	79	52	93	34	0	
3.00	0	1	30	26	19	6	0	
3.25	0	1	7	25	23	66	4	
3.50	0	0	2	8	7	48	17	
3.75	2.25	2.50	2.75	3.00	3.25	3.50	3.75	

(c) MobileNet-V2

True Class	2.25	2.50	2.75	3.00	3.25	3.50	3.75	
Predicted Class	2.25	112	90	11	0	0	1	0
2.25	76	574	182	9	6	1	1	
2.50	15	461	305	33	43	10	5	
2.75	0	18	133	46	58	14	10	
3.00	0	7	32	16	19	4	4	
3.25	0	3	6	12	32	34	39	
3.50	0	2	3	5	2	33	37	
3.75	2.25	2.50	2.75	3.00	3.25	3.50	3.75	

(d) EfficientNet

Fig. 15. Confusion matrix of four network models.

shown in Fig. 16. The accuracies of XceptionNet and LetNet-5 were 88.7 % and 88.22 %, respectively, but the detection efficiency of the latter was much higher than that of the former. Compared with XceptionNet, its accuracy was improved by 1.32 %, and the recall and F1 value were reduced by 1.66 % and 0.84 %, respectively. The slight fluctuation of the indicators was mainly related to the difference in the learning ability of the two types of networks from class 3.00 to class 3.75. Among the four algorithms, due to the superiority of EfficientNet and MobileNet-V2 under extreme BCS values (2.25 and 3.75), the accuracy and F1 value of the two algorithms were significantly higher than those of the other two networks by 2.5 to 5 percentage points, and 83 % of the F1 value indicated that the two networks could achieve high precision and recall of body condition inspection and timely detection of the health

management problems of extreme BCS dairy cows so that farmers could take emergency measures to reduce economic losses. Considering the accuracy and recall, EfficientNet had the best performance of 91.21 % and 83.6 %. In particular, the recognition time of a single frame feature image was only 0.0049 s, which was much lower than that of other networks (except LetNet-5). The efficiency and convenience of BCS scoring was not only the core demand in actual production but also the key factor restricting the commercialization of existing research, especially the scoring speed of single frame images by system. In this paper, the times required for image preprocessing, convex hull feature extraction and feature image generation were 0.116 s, 3.00 s and 0.325 s, respectively. Since only one cow was allowed to pass through the indoor channel at a single time, the time interval between the two cow

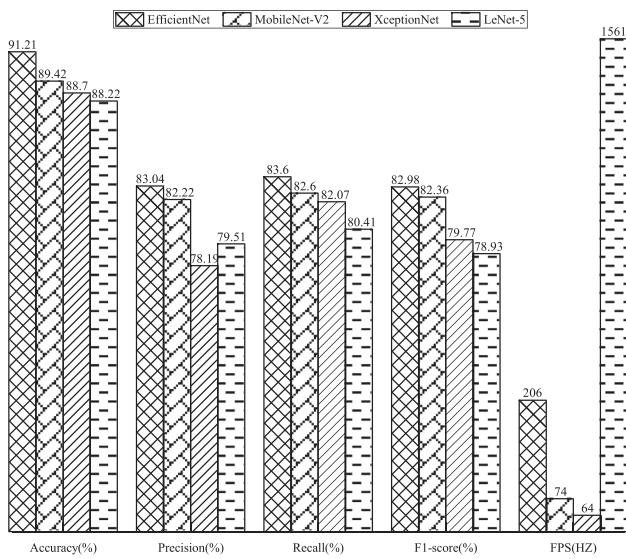


Fig. 16. Recognition results of each network model under an error range of 0.25.

trigger shooting fully meets the time required to generate feature images from the original image. The time taken in the feature extraction stage was 1/60 of the time taken by Fisher et al. (Fischer et al., 2015) to mark key points (3 min), and the system took 3.441 s for a single frame, which was close to the detection level (3 s) when Liu (Liu, 2020) used the first commercial (Lilavar) automatic body condition scoring system. Since the four models all use the same feature image set, it can be seen from the above analysis that the performance order of the four networks is EfficientNet, MobileNet-V2, LeNet-5 and XceptionNet. Although the recognition accuracy of optimal EfficientNet has made breakthrough progress compared with the existing research, it should be noted that the superiority of the model might be related to the fact, as proposed by Ferguson et al., that is, a BCS between 2 and 4 had more obvious characteristics compared with other extreme classes (which have difficulty appearing in actual breeding).

4. Discussion

4.1. Model and abnormal result analysis

This study aimed to improve the automation and accuracy of BCS by extracting the overall three-dimensional sagging feature directly related to the body condition from the back area captured by a single low-cost depth camera and optimizing the classification performance of the model for low-sample classes. On the one hand, in the methods based on feature extraction and model analysis, the use of the overall three-dimensional sagging feature avoided the complexity, redundancy and suboptimal results of the selection of the evaluation area and body condition features in the multi-feature combination method, simplified the cumbersome process of feature design and extraction to reduce the influence of the local weak correlation between the evaluation area and its feature combination and body condition on the final BCS, and realized the evaluation of the overall fat by the external three-dimensional representation of fat enrichment. In addition, the impact of potential errors in a single assessment area on the final BCS value and the difficulty in defining the boundary among regions were also addressed. On the other hand, in methods based on deep learning, compared with direct use of a depth image as the model input, the establishment of convex hull feature image could effectively avoid the interference of the redundant information contained in the forequarters image on the final BCS and eliminated the impact of individual differences, shooting angle, body shaking and other factors of the cow. This improved the

practicability of the system to deal with cows with different BCS values in complex environments. Finally, to further improve the accuracy of the system and solve the precision problem of low sample classes, a two-level classification model was proposed to complete the classification task based on convex hull feature images. The results showed that the EfficientNet network performed best when 2.75 was the classification boundary. The accuracy within the 0.25 errors was 91.21 %, which was 2 % higher than that of the LeNet-5, XceptionNet and MobileNet-V2 networks. The FPS (frames per second) was 206, which met the need for detection effectiveness without human participation.

Although this study achieved high accuracy, the special cases in the model classification results provided some ideas for improving the performance of the model. As shown in Fig. 15, in the range of 0.50 errors, each of the four algorithms had 213 feature images that were accurately identified as real class 2.25, and one was predicted to be 3.50, indicating that the image may be the same. The foreign body in the shooting process or false manual scoring might be the reason for the abnormal judgement of the image. Further analysis showed that the former was because the foreign body was located in the hindquarters of cows and was closely connected with it, which could not be filtered out. The latter might be because cows with a BCS of 3.25 or 3.50 scored 2.25. Since the images were collected in indoor walkways, the reason was that the probability of foreign bodies was lower. In addition, no more than 19 images of the 3.25 category were correctly identified by each of the four networks. However, when the error range was expanded, four networks showed large identification differences, which indicated that the accurately identified image may be a group of images with typical features of class 3.25 and were obviously different from other images of the class. Therefore, the accurate identification rate could be improved by studying the feature difference between this group of images and other images of this kind. It should be noted that addressing the abnormal prediction (error > 0.5) of categories 2.75 and 3.00 was a direct approach to improving the precision of categories 3.50 and 3.75. In summary, the overall performance of the model could be improved in three aspects: excluding special case images with false manual scoring or foreign bodies, studying the feature differences between images within a specific class and improving the anomaly detection of common classes.

4.2. Compared with existing research

The overall accuracies of the model were contrasted against the existing research with medium or high automation levels. In addition, as the BCS score was defined as a classification problem of discrete values, works that compute only metrics suitable for regression analysis and continuous values (such as R, R², RMSE) were not taken into account. Table 4 shows the overall accuracy level of the existing research and the developed system in different error ranges. Compared with other intelligent condition scoring methods based on pattern recognition, both the direct model based on the convex hull feature map and the two-level model had weak accuracy advantages when the error range was 0 or 0.5, but the accuracy within 0.25 errors made breakthrough progress, increasing by 8.22 % and 11.23 %, respectively. The accuracy of convex hull features in describing body condition information and individual differences and the classification superiority of EfficientNet were the keys to improving the accuracy of BCS. The efforts of the secondary model in resolving the low classification accuracy of low sample categories caused by sample imbalance and reducing the potential classification errors of the boundary categories were the reasons for further improvement of the accuracy. As a key index to evaluate the precision of models, the accuracy (91.21 %) within the 0.25 errors was greatly improved, which provided a powerful tool for improving the efficiency and accuracy of body condition assessment in intensive pastures, real-time monitoring of the nutritional health level and production and reproduction performance of cattle, and realizing fine breeding management. The accuracy increased by 5.39 % when the error increased from 0.25 to 0.50, indicating that the two-level model combining convex

Table 4

Overall accuracy level reported by related works and the developed system within different human error ranges.

Error range	Shelly (2016)	Spoliansky (2016)	Rodriguez (2018)	Sun (2019)	Rodriguez (2019)	Direct Model	Our Model
0(exact)	33.80 %	–	40 %	45 %	41 %	45.88 %	45.01 %
0.25	71.35 %	74 %	78 %	77 %	82 %	88.74 %	91.21 %
0.50	93.91 %	91 %	94 %	98 %	97 %	96.33 %	97.60 %

hull features and the EfficientNet network could accurately reproduce the body condition value of dairy cows with a specific BCS (2.25–3.75) and that the error was less than 0.25, which had obvious advantages and met actual needs.

4.3. Future work

Although the body condition detection capability of this method was assessed by using the back depth image set of single breed dairy cows captured in the same dairy farm, its detection ability for different farms, different varieties and extreme body condition (BCS < 2.25 or BCS > 3.75) dairy cows still needs further research and verification. In addition, the system precision can be further improved in three aspects: studying the characteristic differences among classes and within classes to improve the accuracy of individual classes (3.00, 3.25), using multiple models to simultaneously predict to detect exceptional images, and rechecking the abnormal prediction of common classes (2.50, 2.75). Optimization of convex hull feature extraction time and development of mobile terminal systems are keys to achieving efficiency, convenience and low cost. Finally, we should pay attention to the neglected short rib area and further perfect the body condition information of the convex hull feature map.

5. Conclusion

This paper presented a method to extract convex hull features from depth images and applied them to body condition scoring. In this method, the hindquarters extracted from the top-view depth image of the cow during walking were used as the body condition assessment area. Through the feature image formed by the voxelization and convex hull of the hindquarters point cloud, the secondary model based on a convolutional neural network was used to classify it to determine its BCS. The convex hull feature proposed in this paper has a high correlation with the BCS and has good robustness to individual differences. The image proportion with recognition error less than 0.25 of the constructed second-level model based on the EfficientNet network reached 91.21 %. The recognition effect was better than that of the secondary model based on the MobileNet-V2, XceptionNet and LeNet-5 networks, and the average recognition speed was 3.446 frame/s, which remarkably improved the accuracy and efficiency of BCS. Therefore, the method presented in this paper could promote the popularization and application of automatic body condition scoring systems in large-scale dairy farming.

CRediT authorship contribution statement

Kaixuan Zhao: Conceptualization, Methodology, Investigation. **Meng Zhang:** Validation, Writing – review & editing. **Weizheng Shen:** Funding acquisition. **Xiaohang Liu:** Software, Writing – original draft. **Jiangtao Ji:** Supervision. **Baisheng Dai:** Validation, Formal analysis, Writing – review & editing. **Ruihong Zhang:** Formal analysis, Writing – review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The authors do not have permission to share data.

Acknowledgements

The authors acknowledge University of Kentucky for facilitation of data acquisition and permission for data use. This study was supported by the National Key R&D Plan Key projects of Scientific and technological Innovation Cooperation between Governments (Grant No. 2019YFE0125600), and the National Natural Science Foundation of China (Grant No. 32002227, 32072788, 31902210).

References

- Albornoz, R.I., Giri, K., Hannah, M.C., Wales, W.J., 2022. An improved approach to automated measurement of body condition score in dairy cows using a three-dimensional camera system. *Animals* 12, 72.
- Azzaro, G., Caccamo, M., Ferguson, J.D., Battiatto, S., Farinella, G.M., Guarnera, G.C., Puglisi, G., Petriglieri, R., Licitira, G., 2011. Objective estimation of body condition score by modeling cow body shape from digital images. *J. Dairy Sci.* 94 (4), 2126–2137.
- Barber, C.B., Dobkin, D.P., Huhdanpaa, H., 1996. The quickhull algorithm for convex hulls. *ACM Transactions on mathematical software (TOMS)*. 22 (4), 469–483.
- Bercovich, A., Edan, Y., Alchanatis, V., Moallem, U., Parmet, Y., Honig, H., Maltz, E., Antler, A., Halachmi, I., 2013. Development of an automatic cow body condition scoring using body shape signature and fourier descriptors. *J. Dairy Sci.* 96 (12), 8047–8059.
- Bewley, J.M., Peacock, A.M., Lewis, O., Boyce, R.E., Roberts, D.J., Coffey, M.P., Kenyon, S.J., Schutz, M.M., 2008. Potential for estimation of body condition scores in dairy cattle from digital images. *J. Dairy Sci.* 91 (9), 3439–3453.
- Chebel, R.C., Mendonça, L.G.D., Baruselli, P.S., 2018. Association between body condition score change during the dry period and postpartum health and performance. *J. Dairy Sci.* 101 (5), 4595–4614.
- Fischer, A., Luginbühl, T., Delattre, L., Delouard, J.M., Faverdin, P., 2015. Rear shape in 3 dimensions summarized by principal component analysis is a good predictor of body condition score in Holstein dairy cows. *J. Dairy Sci.* 98 (7), 4465–4476.
- Huang, X., 2020. Individual information perception and body condition score assessment for dairy cow based on multi-sensor. University of Science and Technology of China, Hefei.
- Huang, X., Hu, Z., Wang, X., Yang, X., Zhang, J., Shi, D., 2019. An improved single shot multibox detector method applied in body condition score for dairy cows. *Animals* 9, 470.
- Imamura, S., Zin, T.T., Kobayashi, I., Horii, Y., 2017. Automatic evaluation of cow's body-condition-score using 3D camera. 2017 IEEE 6th Global Conference on Consumer Electronics (GCCE).
- Kristensen, E., Dueholm, L., Vink, D., Andersen, J.E., Jakobsen, E.B., Illum-Nielsen, S., Petersen, F.A., Enevoldsen, C., 2006. Within- and Across-Person Uniformity of Body Condition Scoring in Danish Holstein Cattle. *J. Dairy Sci.* 89 (9), 3721–3728.
- Liu, Y., 2020. A comparative study of automatic body condition scoring and manual scoring. *China Dairy*. 224, 55–60.
- Martin, R.A., Ehle, F.R., 1986. Body composition of lactating and dry holstein cows estimated by deuterium dilution. *J. Dairy Sci.* 69 (1), 88–98.
- Mirjalili, S., Lewis, A., 2016. The whale optimization algorithm. *Adv. Eng. Softw.* 95, 51–67.
- Petrovska, S., Jonkus, D., 2014. Relationship between body condition score, milk productivity and live weight of dairy cows. *Int. Sci. Conf. Proc.* 1, 100–106.
- Rathbun, F.M., Pralle, R.S., Bertics, S.J., Armentano, L.E., Cho, K., Do, C., Weigel, K.A., White, H.M., 2017. Relationships between body condition score change, prior mid-lactation phenotypic residual feed intake, and hyperketonemia onset in transition dairy cows. *J. Dairy Sci.* 100 (5), 3685–3696.
- Roche, J.R., Dillon, P.G., Stockdale, C.R., Baumgard, L.H., VanBaale, M.J., 2004. Relationships among international body condition scoring systems. *J. Dairy Sci.* 87 (9), 3076–3079.
- Rodríguez, A.J., Arroqui, M., Mangudo, P., Toloza, J., Jatip, D., Rodríguez, J.M., Teyseyre, A., Sanz, C., Zunino, A., Machado, C., Mateos, C., 2018. Body condition estimation on cows from depth images using Convolutional Neural Networks. *Comput. Electron. Agric.* 155, 12–22.
- Rodríguez, A.J., Arroqui, M., Mangudo, P., Toloza, J., Jatip, D., Rodríguez, J.M., Teyseyre, A., Sanz, C., Zunino, A., Machado, C., Mateos, C., 2019. Estimating body

- condition score in dairy cows from depth images using convolutional neural networks, transfer learning and model ensembling techniques. *Agronomy* 9, 90.
- Shelly, A.N., 2016. Incorporating machine vision in precision dairy farming technologies. University of Kentucky, Lexington.
- Song, X., Bokkers E, E.A.M., Mourik, S.V., Koerkamp, P.W.G.G., Tol, P.P.J.V.D., 2019. Automated body condition scoring of dairy cows using 3-dimensional feature extraction from multiple body regions. *J. Dairy Sci.* 102, 4294–4308.
- Spoliansky, R., Edan, Y., Parmet, Y., Halachmi, I., 2016. Development of automatic body condition scoring using a low-cost 3-dimensional Kinect camera. *J. Dairy Sci.* 99, 7714–7725.
- Sun, Y., Huo, P., Wang, Y., Cui, Z., Li, Y., Dai, B., Li, R., Zhang, Y., 2019. Automatic monitoring system for individual dairy cows based on a deep learning framework that provides identification via body parts and estimation of body condition score. *J. Dairy Sci.* 102, 10140–10151.
- Tan, M., Le, Q.V., 2019. EfficientNet: rethinking model scaling for convolutional neural networks. Proceedings of the 36th International Conference on Machine Learning.
- Wu, F., Wei, J., Zhang, M., 2013. Kernel-PCA application on dairy cow automatic body condition score. *Optical Technique.* 39, 222–227.
- Zhao, K., 2017. Dairy cattle's information perception and behavior analysis based on machine vision. Northwest A&F University, Yangling.
- Zhao, K., Shelley, A.N., Lau, D.L., Dolecheck, K.A., Bewley, J.M., 2020. Automatic body condition scoring system for dairy cows based on depth-image analysis. *Int. J. Agric. & Biol. Eng.* 13, 45–54.
- Zin, T.T., Seint, P.T., Tin, P., Horii, Y., Kobayashi, I., 2020. Body condition score estimation based on regression analysis using a 3D camera. *Sensors* 20, 3705.