

Research Paper

Application of weight prediction for Holstein dairy cows in non-pregnant and postpartum stages

Hsin-I Chiang ^{a,c} , Jia-Ming Zhou ^b , Wen-Lin Chu ^{b,c,*} ^a Department of Animal Science, National Chung-Hsing University, Taichung, 402202, Taiwan^b Department of Bio-industrial Mechatronics Engineering, National Chung-Hsing University, Taichung, 402202, Taiwan^c Smart Sustainable New Agriculture Research Center (SMARTer), Taichung 40227, Taiwan

ARTICLE INFO

ABSTRACT

Keywords:

Postpartum stages
Body weight prediction
Health monitoring
Machine learning

A non-contact weight prediction system for Holstein dairy cows was developed based on depth sensing technology, designed to predict weight changes during non-pregnant and postpartum stages. The system utilises an Intel RealSense D455 depth camera to capture depth image information from cow's dorsal, hips, and side regions, extracting effective body surface feature data through a systematic data processing workflow. Experimental results demonstrate that the Gaussian Process Regression (GPR) model performed most excellently in the cow's dorsal region. For example, with cow number cid603 during the non-pregnant period, prediction accuracy reached a root mean square error (RMSE) of 19.37 kg and a mean absolute percentage error (MAPE) of 1.82 %; with cow number cid700 in the postpartum stage, the model maintained an RMSE of 22.35 kg and MAPE of 2.74 %, exhibiting robust model generalisation capability. Compared to traditional farm methods based on body length and heart girth measurements, the weight prediction system proposed in this study significantly improved the accuracy and stability of weight prediction, especially in capturing physiological state changes (such as postpartum weight loss). Experimental results indicate that the GPR model exhibited the best predictive ability and generalisation with feature data from the dorsal region, effectively supporting precise monitoring of dairy cow weight. Future research directions should focus on optimising image preprocessing techniques, incorporating more physiological parameters (such as feed intake), and integrating depth information from different angles to enhance the system's adaptability in complex environments, thereby strengthening the universality and reliability of the weight prediction model.

Nomenclature table

(continued)

Abbreviations/ Symbols	Description
ARD	Automatic Relevance Determination
AUC	Area Under Curve
<i>a</i>	Exponential regression model parameter
BCS	Body Condition Score
BG	Belly Girth
BL	Body Length
BW	Body Weight
<i>b</i>	Exponential regression model parameter
$b_j^{(1)}$	Bias value of hidden layer node <i>j</i>
$b_h^{(2)}$	Bias value of output layer node <i>h</i>
CNN	Convolutional Neural Network
<i>c</i>	Double-exponential regression model parameter

(continued on next column)

<i>d</i>	Double-exponential regression model parameter
D ₁	Distance from side camera to ground (cm)
D ₂	Distance from dorsal camera to ground (cm)
D ₃	Distance from hips camera to ground (cm)
D ₄	Distance from crossbeam under farm eaves to ground (cm)
D ₅	Distance from farm eaves to ground (cm)
FNN	Feedforward Neural Network
f ₁	"LeakyReLU" activation function of the hidden layer
f ₂	Linear function of the output layer
<i>f(x)</i>	Ratio of pixels to area function
GPR	Gaussian Process Regression
HG	Heart Girth
IoU	Intersection over Union
<i>k</i>	Number of subsets in k-fold cross validation
MAE	Mean Absolute Error (kg)
MAPE	Mean Absolute Percentage Error (%)

(continued on next page)

* Corresponding author. Department of Bio-industrial Mechatronics Engineering, National Chung-Hsing University, Taichung, 402202, Taiwan.

E-mail address: wilchu@nchu.edu.tw (W.-L. Chu).

(continued)

MSE	Mean Squared Error (kg^2)
MSE_l	Mean Squared Error of the l -th iteration (kg^2)
MSE_{avg}	Average Mean Squared Error (kg^2)
m	Number of neurons in hidden layer
NLS	Nonlinear Least Squares
N	Total number of samples
n	Number of input features in FNN
O_1	First detected target object
O_2	Second detected target object
ROC	Receiver Operating Characteristic curve
R^2	Coefficient of determination
RMSE	Root Mean Square Error (kg)
WH	Withers Height
$w_{ji}^{(1)}$	Weight from input layer node i to hidden layer node j
$w_{hj}^{(2)}$	Weight from hidden layer node j to output layer node h
x	Distance from camera to cow (cm)
x_i	The i -th element of input vector
y	Actual weight of sample (kg)
\hat{y}	Predicted weight of sample (kg)
z	Total number of prediction outputs
θ_1	Side camera capture angle (Degree)
θ_2	Dorsal camera capture angle (Degree)
θ_3	Hips camera capture angle (Degree)
Δ^2	Second-order Laplacian operator

1. Introduction

This research aims to enhance the efficiency and precision of dairy cow management in modern livestock farming, addressing challenges faced in traditional dairy cow management models with high labour dependency, rising labour costs, and labour shortages (Sharpe & Heins, 2023). Currently, many dairy farms are confronted with substantial challenges in daily management and animal care due to a shortage of human resources. In particular, critical tasks such as body weight monitoring, health assessment, body condition scoring (BCS), and precision feed management each requiring specialised knowledge and experiential judgment often suffer from insufficient operational efficiency and compromised assessment accuracy. These limitations may indirectly undermine the health maintenance and productive performance of dairy cow. Traditionally, on-site health evaluations in livestock farming have relied heavily on subjective judgment based on the experience of field personnel (Heinrichs et al., 2017; Martins et al., 2020). Such practices tend to introduce assessment bias and impede the precise acquisition of vital indicators such as body weight, thereby reducing the reliability and scientific rigor of managerial decision-making processes. For example, in early postpartum, dairy cows with more significant weight loss show significantly reduced embryo quality and pregnant rates (Carvalho et al., 2014); in late pregnant, dairy cows with higher body condition scores (BCS) experience more significant weight loss, accompanied by stronger lipolysis and insulin resistance, thereby increasing the risk of metabolic diseases (Rico et al., 2015). These findings highlight the importance of monitoring and managing dairy cow weight changes, which may influence overall health and production efficiency.

In the current situation where smart livestock farms are not widespread, traditional farms primarily estimate dairy cow weight through two methods. First, by using a tape measure to estimate weight by measuring body length (BL) and heart girth (HG), a method initially proposed by Milner and Hewitt (1969) for estimating horse weight; subsequently, Carroll and Huntington (1988) evaluated and improved this formula and further defined horse body length (Martinson et al., 2014). Wagner and Tyler (2011) research also compared various methods for estimating horse weight. This formula was gradually applied as a basis for estimating cow weight (Silva et al., 2024). Secondly, manual measurement of dairy cow is often employed, followed by statistical analysis and the estimation of body weight (BW) through the fitting of regression models (Enevoldsen & Kristensen, 1997). When manually measuring health information such as withers height (WH),

heart girth (HG), body length (BL), and belly girth (BG), to ensure measurement accuracy (Erickson et al., 2020; Heinrichs & Hargrove, 1994; Rotondo et al., 2021), it is necessary to ensure that the cows are in a natural standing position, avoid any feeling of compression, maintain the head facing forward, and ensure appropriate tape measure tension. In experimental studies of various cow body parts, results indicate that heart girth (HG) is the single most effective weight prediction variable, showing a high positive correlation with weight (Pearson correlation coefficient of 0.88) (Heinrichs et al., 2007; Yan et al., 2009). For example, Heinrichs et al. (2007) developed a weight formula applicable to Holstein dairy heifers, which is less affected by variations between different observers (standard deviation of heart girth measurements about 2–3 cm), and most estimation errors are below 8 %, particularly suitable for farm environments with relatively limited resources. Yan et al. (2009) further verified that combining heart girth (HG) with belly girth (BG) and body length (BL) can significantly improve estimation accuracy, with a coefficient of determination of $R^2 = 0.78$.

Although the above methods have made particular progress in cow weight estimation, manual measurement methods still have several inherent limitations, including insufficient measurement repeatability, time-consuming and labour-intensive processes, potential stress to cow, and subjective judgments due to differences in measuring personnel's operational methods, thereby increasing uncertainty in health information (Zhao et al., 2023). In traditional measurement methods, inspection personnel's subjectivity and experience differences easily lead to variations in cow health information. The influence of subjective factors results in poor repeatability of health information, and even the same inspection personnel may have different assessment results when evaluating the same dairy cow at different times. Furthermore, manual measurements require considerable time and manpower, especially in large farms where many dairy cows need to be evaluated, making the assessment efficiency relatively low and difficult to meet the needs of modern livestock industry development (Martins et al., 2020; Rodríguez Alvarez et al., 2018).

With the rapid development of machine vision technology, in response to the trend of agricultural and livestock modernization transformation in various countries, many scholars are actively engaged in developing new technologies for cow health management based on the needs of smart farms. Multiple studies have demonstrated the superior performance of machine learning methods in animal weight prediction applications. Preethi et al. (2023) applied artificial neural networks (ANN) and nonlinear regression models for weight prediction of Landry pigs across different age stages, with results showing that ANN models achieved higher prediction accuracy compared to traditional regression methods. Ruchay et al. (2022) conducted a comprehensive comparative analysis of machine learning algorithms for Hereford cow weight prediction, finding that the ExtraTreesRegressor algorithm demonstrated the best performance among all tested machine learning algorithms. Furthermore, Ruchay et al. (2022c) proposed a cow weight prediction method based on RGB-D image depth regression, achieving 91.6 % measurement accuracy on real datasets. Non-contact sensing technology demonstrates significant advantages in effectively overcoming the limitations and deficiencies of traditional manual measurements. This technology can collect weight and health information through cameras during dairy cow's daily activities, avoiding the cumbersome process of traditional manual measurements and effectively reducing staff workload and the stress caused to dairy cows (Wang et al., 2021). According to different types of image data, Lu et al. (2022) proposed a robust livestock animal pose standardization method based on 2D and 3D fusion. This approach integrates data from multiple RGB-D cameras and, within relatively controlled environments, effectively addresses pose variation issues in 3D point cloud data. Le Cozler et al. (2019) developed and validated a high-precision three-dimensional scanning system called Morpho3D, which combines dairy cow morphological features (such as heart girth, chest depth, hips width, etc.) with weight prediction. This system can precisely capture their

body contours and estimate weight by generating a complete three-dimensional point cloud map of dairy cow. Compared to traditional manual measurement methods, the Morpho3D system demonstrates a higher correlation in weight accuracy, with Pearson correlation coefficients of $R = 0.89$ and $R = 0.76$ between chest depth and body width, respectively, and the target variable. Martins et al. (2020) used Microsoft Kinect V2 depth cameras to photograph the backs and sides of 28 Holstein dairy cow and 27 Holstein heifers, establishing fitting models based on three-dimensional images. Experimental results showed that the weight prediction model based on dorsal images had a coefficient of determination of $R^2 = 0.96$. In contrast, the weight prediction model based on side images had a coefficient of determination of $R^2 = 0.89$. Dohmen et al. (2021) proposed an automatic dairy cow weight prediction model based on Mask R-CNN segmentation method and Convolutional Neural Network (CNN). The study used Microsoft HD-3000 LifeCam webcams to collect visible light images of dairy cow sides and backs. It used deep learning technology to automatically segment the visible light images of dairy cow, extract features, and estimate weight. Experimental results showed that the model's prediction evaluation indicators for side images and dorsal images were coefficient of determination $R^2 = 0.91$ and $R^2 = 0.96$, and root mean square error RMSE = 27 kg and RMSE = 20 kg, respectively.

In summary, as the modern livestock industry accelerates toward intelligent development, traditional manual methods for measuring cattle weight and health information are gradually being replaced by non-contact technologies due to their high labour requirements, insufficient repeatability, and potential stress effects on cattle. Research findings consistently demonstrate that cattle weight and health management methods based on non-contact technologies significantly outperform traditional manual measurement results in terms of Pearson correlation coefficients and coefficients of determination, thereby establishing the feasibility of implementing smart farms in the future.

Based on this development trend, this study addresses deficiencies in existing research by proposing innovative solutions, focusing on significant breakthroughs in cattle weight prediction technology. Specifically, this research objectives are: (1) to develop a physiological state-specific weight prediction system designed explicitly for non-pregnant and postpartum stages; (2) to systematically evaluate prediction accuracy across three body regions—dorsal, hips, and side areas—using Intel RealSense D455 depth-sensing cameras; and (3) to provide effective camera positioning solutions for diverse farm configuration requirements through comparative analysis of machine learning

approaches.

This study employs a comparative analysis methodology between Feedforward Neural Networks (FNN) and Gaussian Process Regression (GPR) to systematically evaluate predictive performance across different physiological states. To ensure robustness in practical applications, the system integrates a ResNet 50-based automated image quality filtering mechanism that effectively removes cattle movement artifacts and environmental interference factors. Through a 16-week longitudinal validation experiment spanning different physiological states, the results demonstrate that GPR exhibits superior generalization capability in postpartum monitoring scenarios and confirm the practical feasibility of the system under real farm conditions.

2. Materials and methods

To develop a robust non-contact image weight prediction system, this study systematically conducted experimental design and data analysis based on the experimental process shown in Fig. 1 as the core architecture. This study used depth cameras to collect image data of the experimental dairy cow's side (No.), dorsal (No.), and hips (No.) and processed the depth information for each body part. The data processing includes cow recognition, missing value imputation, distance filtering, morphological operations, and connected component analysis. This is to obtain more complete dairy cow surface depth information and effectively filter out unnecessary depth noise. Subsequently, effective and high-quality images of each body part were screened through an image classifier, and area fitting models for each body part were established using the area of the calibration board. Other feature data included dairy cow height, the pixel count of each body part, and the distance between each cow body part and the camera (taking average value and median). Finally, this study used Feedforward Neural Network (FNN) and Gaussian Process Regression (GPR) models for weight prediction and evaluated the prediction results. Detailed experimental details will be elaborated in subsequent sections.

2.1. Data collection

2.1.1. Location and animals

The experimental site for this study is located at the practice farm of the Department of Animal Science at National Chung Hsing University in Taiwan. The farm primarily serves as a practice site for Animal Science students, and five Holstein dairy cows (numbered cid111, cid154,

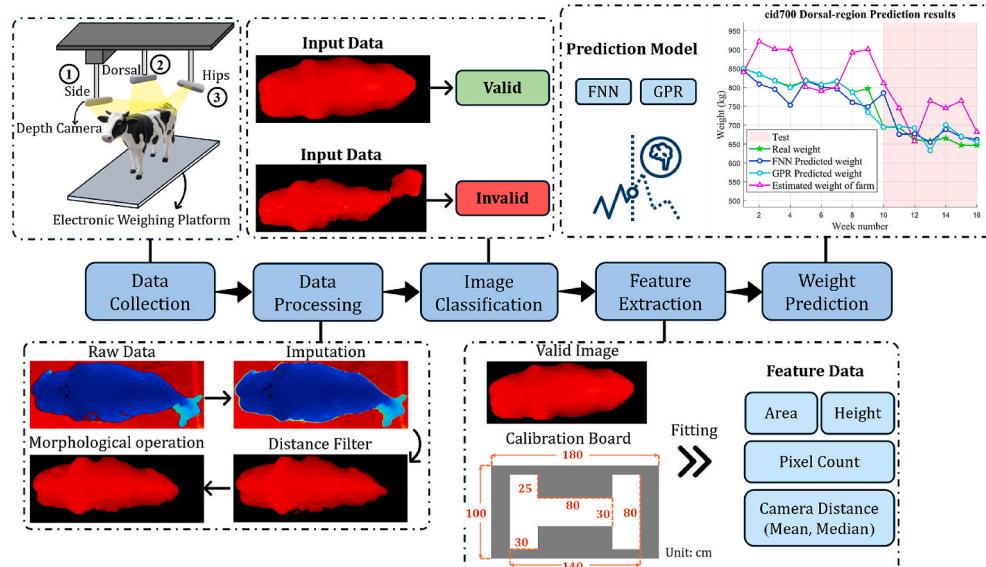


Fig. 1. Flow chart of the main experimental items in this study.

cid603, cid660, and cid700) kept at the farm were the subjects of this experiment. This study was conducted in conjunction with the interdisciplinary, integrated research project of The ENABLE Center at National Chung Hsing University (Project title: “Development of an Intelligent Dairy Cow Management System Based on Objective Body Condition Assessment and Automated Image Recognition”). To accommodate routine farm operations, dairy cow weight measurements were scheduled once per week. Data collection for this study began at 7 a.m. every Wednesday and continued according to the project schedule from June 12, 2024, to December 11, 2024, for 16 weeks. Each dairy cow had 1920 depth images captured from three angles (dorsal, hips, and side regions) during this period. Considering that during the data collection period, cows numbered cid111 and cid700 both experienced pregnant periods and calving stages, causing significant weight fluctuations, this verified the system model’s robustness and repeatability. The remaining cows did not enter pregnant periods and had no calving-related records.

2.1.2. Hardware system setup

For hardware system setup, this study used the Intel® RealSense™ Depth Camera D455 as the main equipment for data collection, with a depth image resolution of 848×480 . The data collection location was set in the electronic weighing platform shown in the site map in Fig. 2. During the data collection process, to avoid direct contact with the cow, farm personnel guided the cow into the electronic weighing platform using feed, facilitating image capture during the cow weighing operation. For detailed explanations of the image capture methods, please refer to Section 2.1.3.

This study meticulously tested camera installation positions, distance from the ground (D), and angles (θ) to obtain high-quality dairy cow images. During the testing period, the research team comprehensively evaluated multiple environmental factors, including environmental light intensity, weather conditions, and potential disturbances

such as cow licking. Given that depth cameras primarily rely on structured light or active infrared light projection principles for depth sensing, stronger environmental light conditions may reduce the accuracy of depth sensors (Hatakeyama et al., 2023; Li et al., 2023). Therefore, this study installed cameras under the eaves above the electronic weighing platform, supplemented with extended support rods to reduce environmental light intensity effectively, as shown in Fig. 3.

In Fig. 3, to ensure effective capture of the cow, the ground distance (D) and angle (θ) for the depth cameras at each position were adjusted. The ground distances for the side, dorsal, and hips cameras were $D_1 = 212$ cm, $D_2 = 195.7$ cm, and $D_3 = 276.8$ cm, respectively, with capture angles of $\theta_1 = 50^\circ$, $\theta_2 = 90^\circ$, and $\theta_3 = 40^\circ$. The distances from the ground (D) for the crossbeam under the farm eaves and the farm eaves were $D_4 = 327$ cm and $D_5 = 382$ cm, respectively. Through the above hardware setup, this study divided the electronic weighing platform dimensions in Fig. 4 into coordinate points. This serves as a quantitative basis for removing the depth image background (see Section 2.2) and estimating the area features of various dairy cow body parts (see Section 2.4).

2.1.3. Shooting method

Before the daily shooting operation began (before 7 a.m.), the farm routinely performed feeding, milking, and measurement records of health information such as withers height (WH), heart girth (HG), and body length (BL) for the experimental cow. To reduce weight fluctuations caused by feeding, milking, drinking water, and environmental factors, the farm had established standardised measurement conditions, including fixed feeding and measurement periods and appropriate water restriction measures before measurement, to enhance data reliability and consistency (Kuzuhara et al., 2015; Wangchuk et al., 2018).

To reduce unnecessary disturbances to dairy cows during their movement and to accommodate farm operational processes, this study formulated corresponding shooting methods and operational procedures based on the movement directions marked in Fig. 2, as shown in Table 1. To avoid image blurring or data loss, the research selected a solid-state drive with a writing speed of 1000 MB/s (model: INTEL SSDPEKNW512G8) and a USB 3.1 standard interface to ensure stable and efficient data transmission.

2.2. Data processing method

During the data collection process, as the depth camera can simultaneously capture visible light, infrared, and depth images (including raw depth data), this study selected the original depth data (raw format) for subsequent data processing and analysis. This refers to the unprocessed output data from the depth sensor (resolution: 848×480 , 16-bit depth) to maximally preserve its original bit depth and sensor noise characteristics (Carfagni et al., 2017).

After obtaining the image data, this study extracted feature data for each body part through the systematic data processing method shown in Fig. 5. First, to correct the viewing angle bias in the original depth data, this study performed rotation correction processing. This was necessary because, as Section 2.1.2 mentioned, environmental factors made it difficult to maintain a direct frontal viewing angle when setting up the cameras. Therefore, this study adopted object detection technology based on the YOLOX model (Ge et al., 2021) to identify cows in the image data collected during the first and second phases. During farm operations, farm personnel routinely guided cows into the electronic weighing platform using the feed, as shown in Fig. 2. However, during this process, cows might deviate from the established path due to surrounding objects or even approach farm workers. As shown in Fig. 6, in the YOLOX model’s recognition results, boundary marker frames might overlap, or the depth information of personnel and cows might be interconnected, making it difficult to effectively capture the body contour boundaries of cow, thereby affecting the subsequent performance of the distance filter, preventing it from precisely filtering out unnecessary part features.

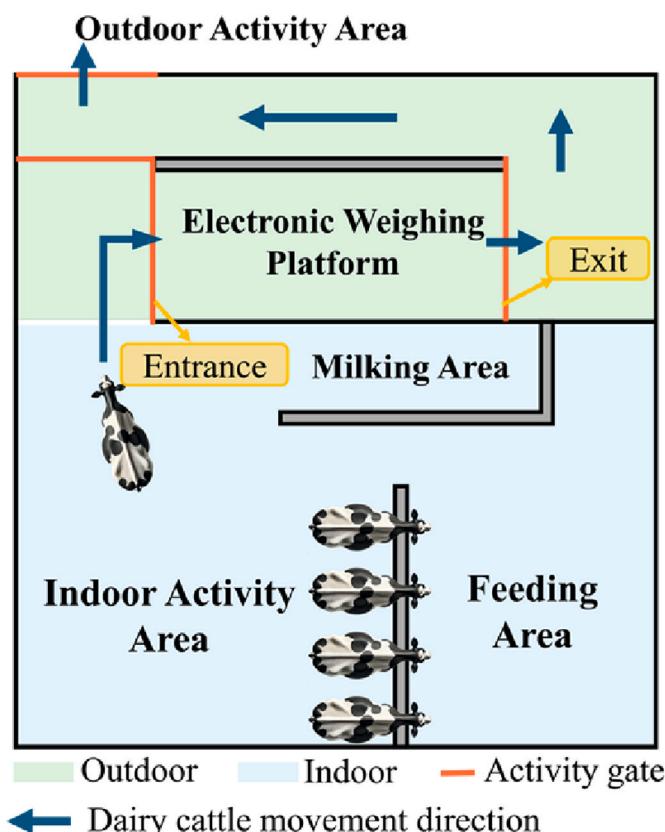


Fig. 2. Schematic diagram of movement direction and purposes across different areas of the dairy farm.

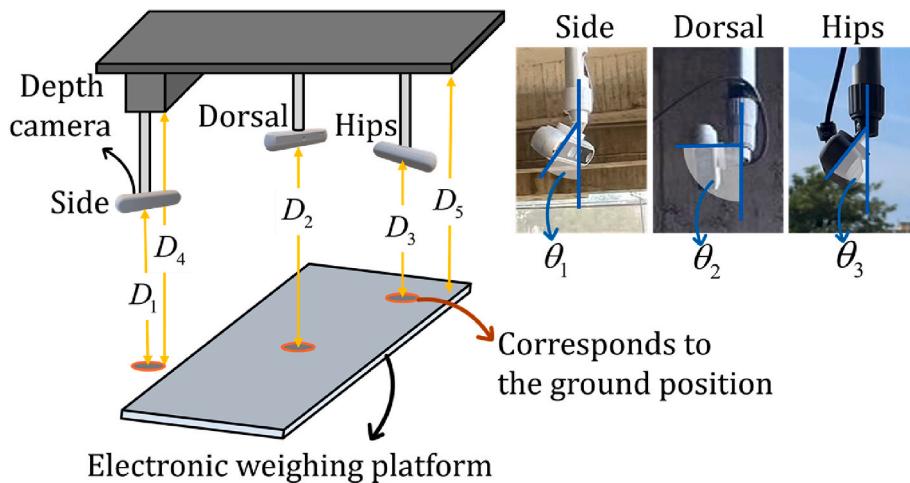


Fig. 3. Installation positions of depth cameras (imaging target areas), distance from the ground (D), angle (θ), and the corresponding vertical coordinate positions on the ground (see Fig. 4).

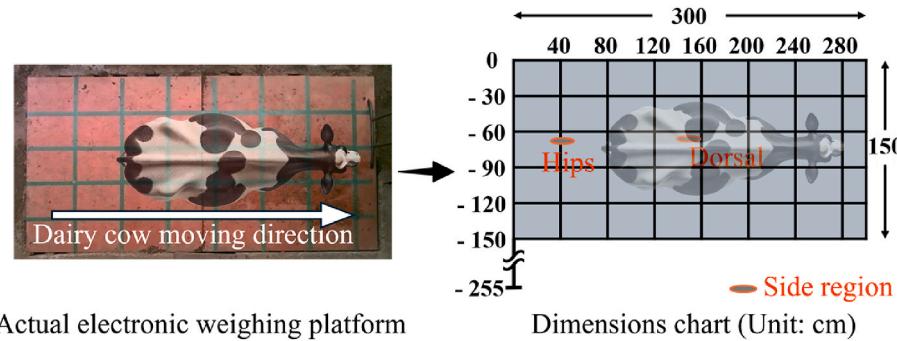


Fig. 4. Schematic diagram of electronic weighing platform scale grid.

Table 1
Shooting operation description.

Shooting Phase	Phase Name	Operation Description
Phase 1	Initial Recording	After the dairy cow enters the designated entrance, the system continuously shoots for 30 s at a rate of 2 fps (A total of 60 images.), while farm personnel measure and record body weight.
Phase 2	Exit Recording	After weight measurement is completed, when the exit gate opens, the system performs the second phase of image collection, and the dairy cow moves to the outdoor activity area along the preset path.

$$\text{IoU} = \frac{|O_1 \cap O_2|}{|O_1 \cup O_2|} \quad (1)$$

To address the problem described, this study introduced Intersection over Union (IoU) analysis to refine the recognition results further. The calculation formula is presented in equation (1). In this Equation, O_1 denotes the first detected target object, while O_2 refers to the second detected target object. The IoU value is calculated as the area of the overlapping region between the boundary boxes of the two target objects, divided by the area of their union. This measure quantifies the degree of overlap between the detected target objects.

To ensure a non-contact state between personnel and cow, a threshold of 0 was established for the IoU value, which means only non-overlapping target object boundary boxes are retained. This approach effectively screens the target objects and enables the cropping of depth information for various parts of the cow's body based on the positions of

the boundary boxes.

Regarding missing value imputation, depth cameras often face sensor technology limitations or physical constraints when capturing transparent, smooth, dark surfaces or distant targets, making it difficult for depth sensors to effectively receive their projected infrared light and obtain valid depth values (Wang et al., 2024). The black and white patches on Holstein dairy cows' body surface easily cause dark areas to absorb infrared light, making it challenging to reflect enough signals and losing depth information. To fill in the missing areas in the depth data, this study used the "inpaint_nans" function provided by MATLAB software for numerical value imputation (D'Errico, 2025). This function is based on a second-order Laplacian operator (Δ^2) least squares method, which reconstructs the depth information of missing areas based on surrounding valid depth values for relatively smooth two-dimensional data.

Subsequently, to further filter out background noise, this study designed corresponding distance filters for each body part based on the characteristics of the depth data. As shown in Fig. 7, the camera position was set as the starting point at 0 cm. The threshold range of the dorsal filter was set from 0 cm to 150 cm, aiming to exclude the ground background and capture dorsal images of cows with heights between 130 cm and 150 cm. The threshold range of the hips filter was set from 0 cm to 100 cm, which was used to exclude the front background and avoid the hips image features becoming too small when cows move forward. The threshold range of the side filter was set from 135 cm to 225 cm to exclude the background behind the cow and avoid situations where cows approach too closely, making it impossible to capture the side area completely (see Fig. 8).

Finally, this study adopted morphological operations and 8-

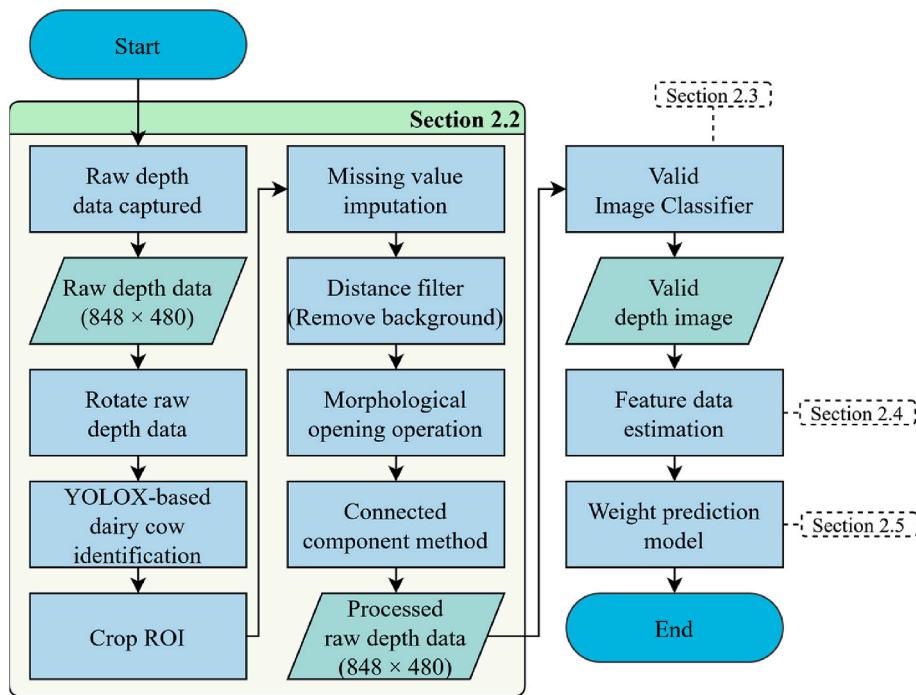


Fig. 5. Integrated analysis process for image data processing and weight prediction.

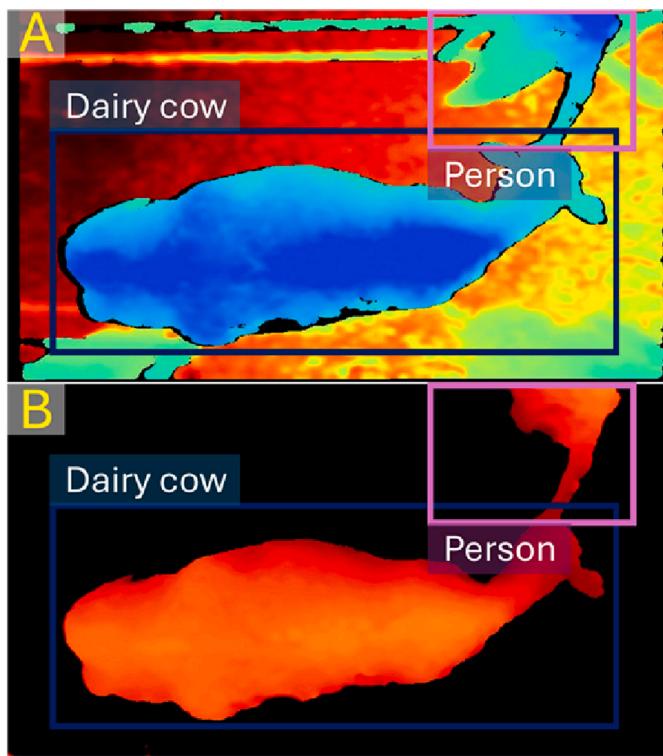


Fig. 6. Examples of occlusion or bounding box overlap caused by personnel during cow guidance process ("A" shows recognition results from the YOLOX model. "B" shows results after distance filtering.).

connected component analysis to filter out noise from body trunk extensions. Morphological operations first performed erosion, followed by dilation, to separate small or slender structures connected to the trunk in most data, such as the head, limbs, and tail. Subsequently, structural regions were screened based on the number of pixels in each connected

component, filtering out components with relatively small pixel counts, thereby preserving the main trunk and parts of the cow in the depth data.

2.3. Valid image classification

During the data collection period, the extended limb structures of cow showed more significant depth variations compared to the core areas (such as dorsal, hips, and side parts). For example, head movement might affect the depth data of the dorsal and side, while tail swinging might affect the depth information of the dorsal, hips, and side. In most cases, the expected noise removal effect could be achieved after the data processing flow described in Section 2.2. However, in a few cases, due to the threshold limitations of the morphological opening operation's structuring elements, the extended structures of the cow body could not be effectively separated, making it impossible for the connected component analysis to filter out such noise precisely. To enhance the precision of the system in the data processing flow, this study introduced a ResNet 50 based effective image classifier in the experimental process shown in Fig. 5, aiming to exclude images containing extra cow body extended structures, as well as images captured when cow entered or exited the electronic weighing platform (i.e., the early images of the first phase and the later images of the second phase). As shown in Tables 1 and if the depth images of various body parts did not contain extra cow body extended structures, the image classifier would determine them as valid images; conversely, if they contained extra cow body extended structures or were images captured when cow entered or exited the electronic weighing platform, as shown in the examples of various body parts' depth images in Table 1, they would be classified as invalid images, with detailed classification results shown in Table 2.

Additionally, to objectively evaluate the performance of the image classification model, this study adopted multiple evaluation indicators, including confusion matrix, verification accuracy, precision, recall, F1 score, receiver operating characteristic curve (ROC curve), and area under curve (AUC). Through these diverse evaluation indicators, the model's performance in distinguishing between valid and invalid images was examined from different perspectives, providing a more comprehensive performance evaluation basis for potential class imbalance

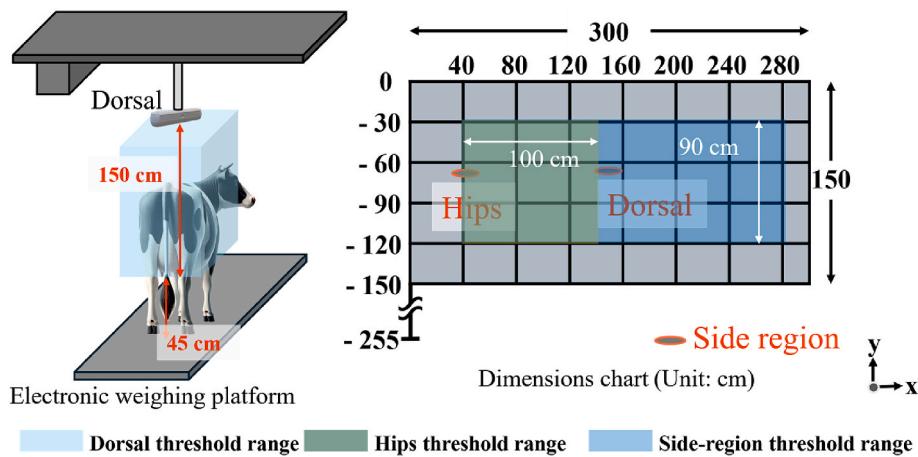


Fig. 7. Distance filter threshold range setting and practical schematic diagram.

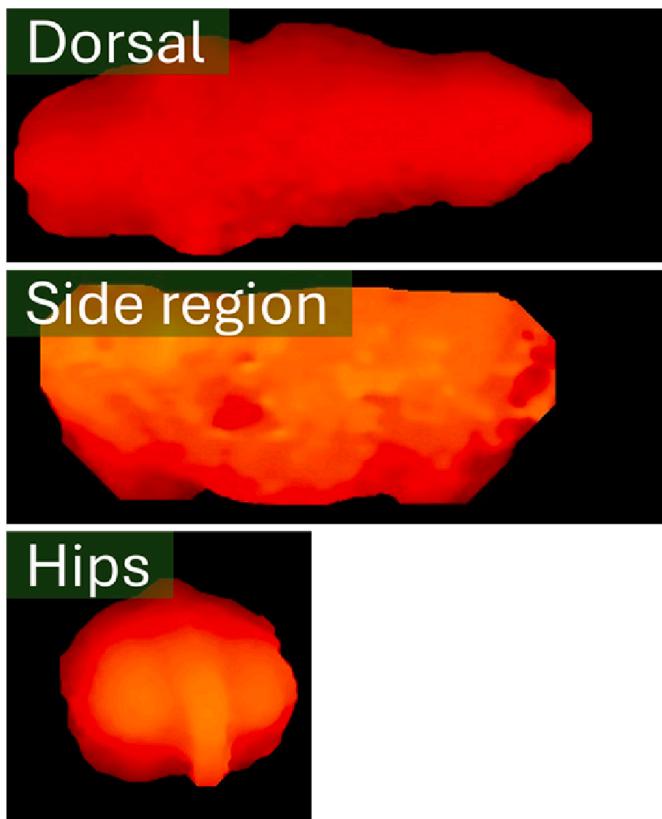


Fig. 8. Example of effective depth image of cow body after processing in Section 2.2.

problems in image data of various body parts that the classification model might face.

2.4. Feature data processing

2.4.1. Calibration board testing

To precisely measure area features of various dairy cow body parts, this study conducted tests using a calibration board based on the coordinates in Fig. 4, simulating area changes of dairy cow at different positions on the electronic weighing platform. The detailed steps of the calibration board testing are shown in Table 3.

Table 2
Detailed description of invalid depth images of cow body.

Image View Type	Description	Corresponding Number in Table 1
Dorsal image	Entering weighing area	a
	Leaving weighing area	b
	Tail overlapping with dorsal	c
	Head raised	d
Side region image	Entering weighing area	e
	Leaving weighing area	f
	With tail movement	g
	With head movement	h
Hips image	Entering weighing area	i
	Leaving weighing area	j
	With tail movement	k

2.4.2. Curve fitting

To extract the feature input data needed for constructing weight prediction models from depth images of various body parts, including area, height, pixel count, and surface depth (i.e., the median and average values of the distance between cow and camera), this section details the processing methods for feature data. For estimating feature area, single-exponential and double-exponential regression provided by MATLAB R2024b Curve Fitting Toolbox were used for data fitting. The double-exponential regression model is shown in equation (2), where x is the distance from the camera to the cow (i.e., calibration board), and $f(x)$ is the ratio of pixels to area. Nonlinear least squares (NLS) were used to estimate the model parameters a , b , c , and d , with the fitting process using the 95 % confidence interval method. The coefficient of determination (R^2) and root mean square error (RMSE) were used to optimise and evaluate the model parameters.

$$f(x) = a \cdot \exp(bx) + c \cdot \exp(dx) \quad (2)$$

This study used the single-exponential regression model shown in equation (3) to fit data for each body part to ensure good data interpretability during the fitting process. Therefore, the best fitting model for each body part will be presented in the experimental results, as detailed in Section 3.2.

$$f(x) = a \cdot \exp(bx) \quad (3)$$

2.5. Weight prediction model

This study used MATLAB R2024b neural network fitting toolbox and statistics and machine learning toolbox to establish feedforward neural network (FNN) and Gaussian process regression (GPR) models for individual cow weight prediction. It compared and evaluated the

Table 3
Calibration board testing experiment.

Experimental Step	Experimental Details
Cow height distribution	Surveyed farm dairy cow height distribution, ranging from 130 cm to 150 cm.
Calibration board configuration	Calibration board covered the expected photography areas of dorsal, side, and hips (as shown in Fig. 10).
Position variable control	Considering that the distance between cow and camera is not fixed, testing was conducted using the coordinate points in Fig. 4, using the calibration board to simulate cow positions and capture images of various body parts, collecting data from multiple positions.
Test parameter design	Test heights were set at 130 cm, 140 cm, 150 cm (as shown in Fig. 11), including hips distances (280 cm, 240 cm, 200 cm, etc.) and side distances (30 cm, 60 cm, 90 cm, etc.), analysing image pixel count changes at various height coordinates to estimate area of body parts.
Calibration board area standard	Actual areas of calibration board: dorsal "T" and side "H" both 7200 cm ² , hips equilateral triangle 1082.5 cm ² (as shown in Fig. 10).
Area estimation method	Established exponential regression models through the ratio relationship between the known actual area of the calibration board and its corresponding image pixel count, along with the average distance data from camera to cow.

performance of these two methods. The input data x refers to the feature data described in Section 2.4.2, and the target data y is individual cow weight.

2.5.1. FNN

Feedforward Neural Network (FNN) effectively learns nonlinear relationships between input and target data when handling complex nonlinear problems (Ojha et al., 2017). Moreover, FNN has a highly flexible structure design, which can be optimally adjusted by tuning the number of network layers, neurons, and activation functions according to the complexity of the data to improve prediction accuracy. Additionally, weights and bias values are automatically adjusted through the error backpropagation mechanism to minimise prediction errors and enhance model performance. The FNN model mainly comprises an input, hidden, and output layer, as shown in equation (4) (Shen et al., 2022).

$$\sum_{h=1}^z \hat{y}_h = f_2 \left(\sum_{j=1}^m w_{hj}^{(2)} f_1 \left(\sum_{i=1}^n w_{ji}^{(1)} \underbrace{x_i}_{\text{Input}} + b_j^{(1)} \right) + b_h^{(2)} \right) \quad (4)$$

In this equation, the input vector $x = \{x_1, x_2, x_3, \dots, x_n\}$ (n is the number of input features), the hidden layer includes $w_{ji}^{(1)}$ which is the weight from input layer node i to hidden layer node j . $b_j^{(1)}$ is the bias value of hidden layer node j . $f_1(\cdot)$ is the "LeakyReLU" activation function of the hidden layer. Then, the output layer includes $w_{hj}^{(2)}$, which is the weight from hidden layer node j to output layer node k . $b_h^{(2)}$ is the bias value of output layer node h . $f_2(\cdot)$ is the linear function of the output layer. Finally, $\sum_{h=1}^z \hat{y}_h$ is the h -th output result, with a total of z outputs.

This study adopted k-fold cross validation for data partitioning and validation to ensure that the model has good generalisation capability. All data were randomly divided into k subsets, with one subset selected as the validation set each time and the remaining $k-1$ subsets combined as the training set. After repeating k times, the average validation error evaluated the model performance, thereby reducing the risk of bias, underfitting, or overfitting caused by data partitioning. Model error used mean squared error (MSE) as the loss function, as shown in equation (5).

$$MSE = \frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2 \quad (5)$$

In this equation, y_i is the actual weight of the i -th sample, \hat{y}_i is the predicted weight, and N is the total number of samples. Finally, the average MSE was calculated through k-fold cross validation, as shown in equation (6), where MSE_l is the mean squared error of the l -th iteration. This method was used to optimise network hyperparameters and improve model generalisation capability.

$$MSE_{avg} = \frac{1}{k} \sum_{l=1}^k MSE_l \quad (6)$$

2.5.2. GPR

This study employed the "fitrgp" function from MATLAB R2024b to construct Gaussian Process Regression (GPR) models. GPR offers the advantage of modelling within a probabilistic framework, providing prediction means and uncertainty assessments (such as confidence intervals). Additionally, GPR can utilise the Automatic Relevance Determination (ARD) mechanism to assign independent length scale parameters to each feature, thereby capturing each variable's importance and influence degree (Fu et al., 2019).

The construction process of the GPR model was as follows.

1. Data Preprocessing: The various input features for cow weight prediction (such as area, height, pixel count, and the median and mean values of the distance between the cow surface and camera) may exhibit nonlinear relationships and varying scales. All input features underwent Z-score standardization in the experiments to ensure consistent feature scaling and enhance model training stability.

2. Defining the Gaussian Process:

a. **Basis Function:** A basis function was selected to capture trends in the data. This study employed the "Constant" basis function to capture constant trends in the data.

b. **Kernel Function:** A kernel function was chosen to define similarity between data points. This study adopted the "ardsquareexponential" kernel function combined with the ARD mechanism to accommodate feature heterogeneity and automatically determine the relevant length scale for each feature.

c. **Noise Standard Deviation (Sigma):** Initial data noise standard deviation values were established.

3. Model Training:

a. The initial covariance matrix between samples was calculated based on the training samples and selected kernel function.

b. Through optimisation algorithms (the toolbox can utilise maximization of Log Marginal Likelihood), hyperparameters (including kernel function parameters and noise standard deviation) were automatically optimised to achieve optimal adjustment of model parameters.

4. Prediction and Evaluation: After model training, the model could be used to predict new weight data and compared with actual weights to evaluate model performance.

5. Model Validation: To ensure the model's generalisation capability under nonlinear relationships, K-fold cross validation was employed to evaluate model performance.

2.5.3. Evaluation metrics

To evaluate the performance of the FNN and GPR weight prediction models, this study employed mean square error (MSE), root mean square

error (RMSE), mean absolute error (MAE), and mean absolute percentage error (MAPE). First, mean square error (MSE) was used to quantify the average squared error between predicted and actual values, as shown in equation (5). Lower MSE values indicate higher prediction accuracy. MSE can directly measure the model's performance in capturing numerical details in cow weight prediction. Next, the root mean square error (RMSE) is shown in equation (7).

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2} \quad (7)$$

In this equation, y_i is the actual weight of the i -th sample, \hat{y}_i is the predicted weight, and N is the total number of samples. Smaller RMSE values indicate more accurate model predictions, while larger RMSE values indicate greater prediction errors. Additionally, to avoid the potential amplification of extreme errors by squared operations, this study comprehensively evaluated using mean absolute error (MAE), where smaller values indicate a more accurate predictive ability of the model, as shown in equation (8).

$$\text{MAE} = \frac{1}{N} \sum_{i=1}^N |y_i - \hat{y}_i| \quad (8)$$

Finally, MAE may lack intuitive comparability when the target variable scales are inconsistent across different datasets or models. Therefore, as shown in equation (9), mean absolute percentage error (MAPE) was adopted to address the interpretability issue of MAE when comparing data of different scales, providing a more intuitive understanding of the proportion of model prediction errors relative to actual values through percentage errors, thereby more clearly evaluating the prediction accuracy of models in different scenarios.

$$\text{MAPE} = \frac{1}{N} \sum_{i=1}^N \left| \frac{(y_i - \hat{y}_i)}{y_i} \right| \times 100\% \quad (9)$$

3. Results and discussion

3.1. Evaluation of classification models for valid and invalid depth images

To ensure depth image quality before subsequent feature estimation, valid image classification was employed for three body regions: Dorsal, Hips, and Side (see Fig. 9). This approach helps avoid bias in subsequent feature extraction results caused by incomplete cow depth images. Following the process in Fig. 5, this study incorporated a binary

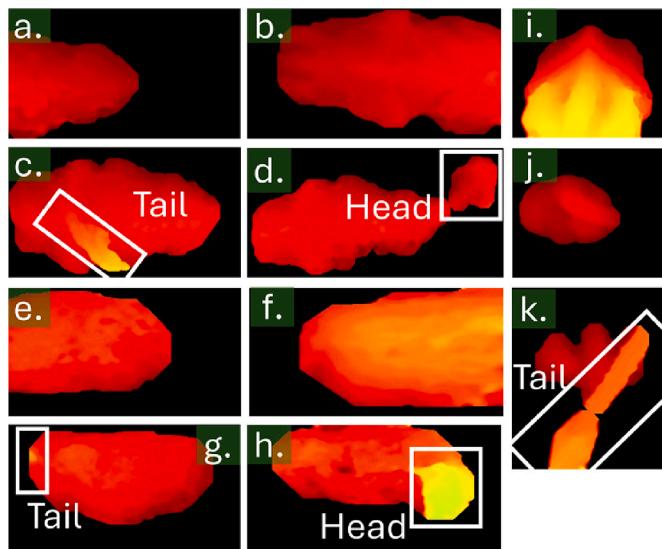


Fig. 9. Example of invalid depth image of cow body after processing in Section 2.2. For details, refer to Table 2.

classification method, categorising images as either "Invalid" or "Valid" to filter out unsuitable depth images. The binary classification results for each body region are shown in the confusion matrices in Fig. 12: (a) The dorsal region classification model demonstrates high true negative (564 cases) and significant true positive performance (343 cases) while maintaining a low misclassification rate, reflecting its robust recognition capability; (b) The hips region classifier exhibits the most outstanding discriminative performance, characterised by extremely low false negatives (only 3 cases) and high true negatives (581 cases), demonstrating its excellence in accurately retaining valid images while effectively excluding invalid samples; (c) Although the side region classification performance shows a relatively higher false positive rate (66 cases), considering the potential occlusion and backlighting factors in depth estimation from side angles, it still maintains acceptable overall classification accuracy. Overall, the classifiers for each body region maintain relatively stable correct judgments in both the "Invalid" and "Valid" categories, indicating that in most cases, the classifiers successfully exclude invalid images and correctly retain useable images.

Table 4 presents quantitative assessment results for a more comprehensive evaluation of the valid image classifier performance, listing overall metrics for each classifier in both "Invalid" and "Valid" categories, including Verify accuracy, Precision, Recall, F1 score, and AUC. Overall, the hips classifier performed best with a verification accuracy 0.9651 across all evaluation metrics. The dorsal classifier achieved a verification accuracy of 0.9380, which, although slightly lower than the hips classifier, is still quite high. The side classifier had a relatively lower verification accuracy of 0.8820. Possible reasons for the decrease in overall classification verification accuracy may be related to unexpected postures of cows during shooting, such as excessive head movements causing overlap with the trunk or turning in place.

Additionally, due to environmental constraints, cow could not be restricted to specific narrow areas for photography. However, the precision, recall, F1 score, and AUC all demonstrate stable and good classification performance. These quantitative evaluation results further validate the effectiveness of the binary classification method, showing that robust classification models can effectively filter out invalid depth images.

3.2. Curve fitting results for cow body area features

This section presents the experimental results of feature estimation methods for each body part according to the analysis process shown in Fig. 5. These results were primarily obtained through extraction from depth images, including cow height, pixel count, and surface depth representing the distance between cow and camera (quantified by both statistical median and arithmetic mean). All features except cow body area estimation could be directly obtained from depth images. Therefore, the following will explain the results of area estimation for various cow body parts. This research design used the ratio of pixel count to area as the input variable and the distance from the calibration board to the camera as the target variable, corresponding to the experimental design presented in Fig. 11. The input data ratio relationships were all calculated from the corresponding pixel counts and known areas obtained from the calibration board's "I" region and "H" region.

In this preprocessing experimental stage, this study conducted single-exponential and double-exponential regression analyses for each body part. To ensure model robustness, a rigorous outlier processing mechanism was implemented during the research process, treating observations falling outside the 95 % confidence interval as outliers and excluding them, thereby enhancing the precision of the fitting process and the reliability of subsequent analysis. The model evaluation employed the coefficient of determination (R^2) and root mean square error (RMSE) to comprehensively assess the applicability of different regression methods for each body part.

The experimental results are shown in Table 5. The dorsal region exhibited the same coefficient of determination ($R^2 = 0.9847$) under

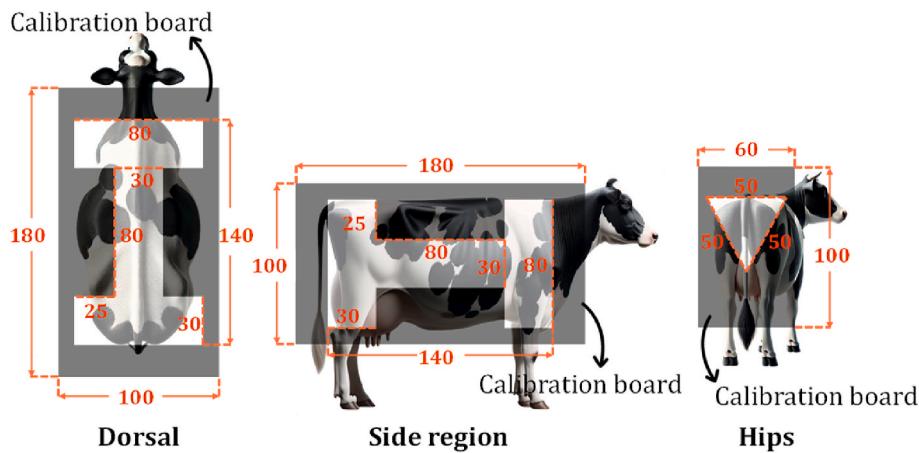


Fig. 10. Schematic diagram of calibration board area testing. (Unit: cm).

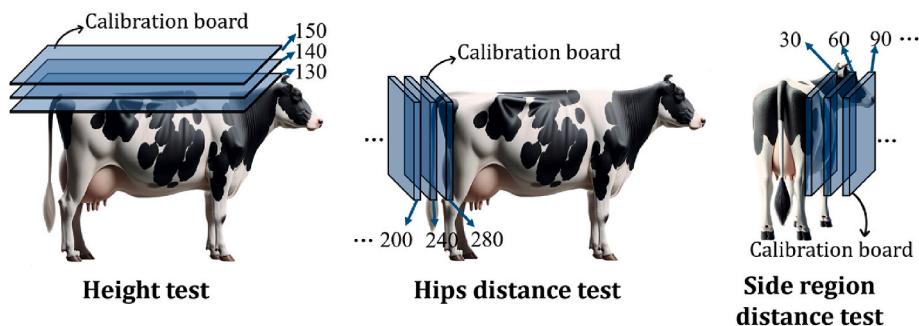
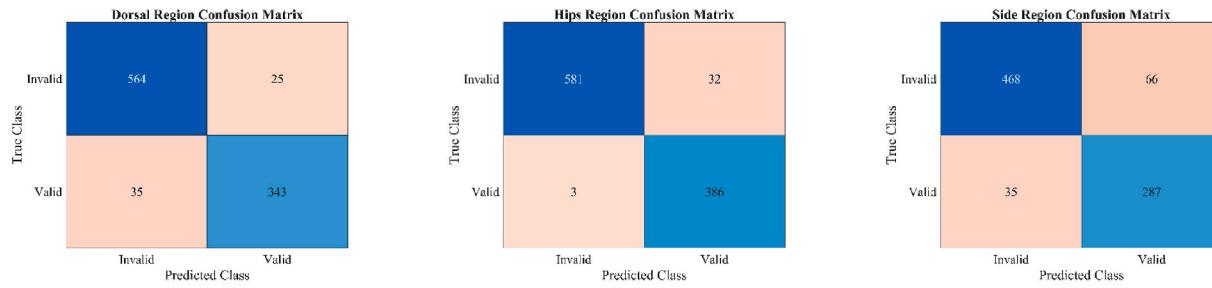


Fig. 11. Schematic diagram of depth testing for various body surface regions. Calibration board positions are tested at each point according to the coordinates in Fig. 4. (Unit: cm).



(a) Dorsal region

(b) Hips region

(c) Side region

Fig. 12. Confusion matrix of effective image classification models for each body region.

single-exponential and double-exponential models. However, the RMSE values showed that the single-exponential model ($RMSE = 0.1186$) had better prediction accuracy than the double-exponential model ($RMSE = 0.1258$), reducing the estimation error by approximately 5.8 %. This result indicates that increasing model complexity did not produce significant fitting advantages in this region. For the hips region, the single-exponential model showed good fit ($R^2 = 0.9811$) but relatively higher error values ($RMSE = 0.3087$); the double-exponential model slightly improved the coefficient of determination ($R^2 = 0.9826$) while simultaneously reducing prediction error ($RMSE = 0.3048$), decreasing the estimation error by about 1.3 %. This result shows that the double-exponential model better fits the hips region. The analysis results for the side region showed that the double-exponential model ($R^2 = 0.9527$, $RMSE = 0.1437$) had significant advantages over the single-exponential model ($R^2 = 0.9478$, $RMSE = 0.1468$), not only improving the

coefficient of determination but also reducing the estimation error by approximately 2.1 %.

Synthesizing the model performance across all body parts, it can be observed that the dorsal region exhibited the highest coefficient of determination under both models, indicating that the planar displacement of cow on the electronic weighing platform has relatively little impact on dorsal depth measurements, providing more stable fitting data. Additionally, the RMSE value of the dorsal region was significantly lower than other parts, further confirming the suitability of this region as a basis for body surface area prediction. The R^2 value of the hips region was slightly lower than the dorsal region, and the RMSE value was significantly higher than other parts. This phenomenon is believed to be related to the morphological changes of the hips and depth variability during cow movement, leading to higher data fluctuation during the fitting process. The double-exponential model demonstrated better

Table 4

Evaluation of effective image classification models.

Region	Category	Precision	Recall	F1 score	AUC	Verify accuracy
Dorsal region	Invalid	0.9416	0.9576	0.9495	0.9795	0.9380
	Valid	0.9231	0.9074	0.9196		
Hips region	Invalid	0.9949	0.9478	0.9708	0.9939	0.9651
	Valid	0.9234	0.9232	0.9566		
Side region	Invalid	0.9304	0.8764	0.9026	0.9507	0.8820
	Valid	0.8130	0.8913	0.8504		

Note: Bold numbers indicate better results.

Table 5

Comparative evaluation of single-exponential and double-exponential models for measuring dairy cow body surface area.

Region	Single-exponential		Double-exponential	
	R ²	RMSE	R ²	RMSE
Dorsal region	0.9847	0.1186	0.9847	0.1258
Hips region	0.9811	0.3087	0.9826	0.3048
Side region	0.9478	0.1468	0.9527	0.1437

Note: Bold numbers indicate better results.

performance in this region, indicating that increasing model complexity can more accurately capture this anatomical region's surface area change characteristics. For the fitting results of the side region, the RMSE value was between the dorsal and hips regions, while the R² value was lower than both dorsal and hips regions. This phenomenon may be due to the oblique characteristics of dairy cow standing posture, requiring more complex mathematical models to describe the surface area changes in the side region accurately. Therefore, the results using the double-exponential model for the side region (R² = 0.9527, RMSE = 0.1437) were significantly better than those using the single-exponential model (R² = 0.9478, RMSE = 0.1468). Finally, all fitting data in this section were obtained from the calibration board test shown in Fig. 10, simulating cow positions on the electronic weighing platform.

3.3. Weight prediction model evaluation and results analysis

This study collected data over 16 weeks as the experimental dataset. The data from the first 9 weeks were used for model training, while data from weeks 10 through 16 were used for model testing. The study employed two methods to establish weight prediction models: Feed-forward Neural Network (FNN) and Gaussian Process Regression (GPR).

Both methods employed k-fold cross-validation during the model training process to ensure robustness and generalisation capability. Through k-fold cross-validation, the training data (weeks 1 through 9) were randomly partitioned into five subsets, with four subsets used for model training each time and the remaining subset serving as the validation set. This approach evaluated model performance under different data partitions and effectively reduced overfitting occurrences (Wong & Yeh, 2020). For out-of-sample testing, data collected from weeks 10 through 16 were utilised for model performance validation. Additionally, all input features underwent z-score standardization processing to ensure feature scale consistency.

The FNN model architecture was configured as follows: The hidden layers adopted a three-layer structure with 50, 25, and 10 neurons, respectively. A dropout mechanism was incorporated with a ratio of 0.2 to avoid overfitting. "LeakyReLU" was selected as the activation function for neurons. L2 regularisation was added to control model complexity with a parameter setting of 0.0001. The batch size was set to 32. The model's optimisation algorithm employed adaptive moment estimation (Adam) to adjust the learning rate dynamically.

The GPR model construction utilised the "fitrgp" function provided by MATLAB R2024b. The basis function was set to "Constant," and the ARD exponential kernel was selected as the kernel function. Model

Table 6

Evaluation indicators of weight prediction models.

Method	MSE	RMSE	MAE	MAPE
FNN	1987.86	44.58	33.33	4.68 %
GPR	2966.95	54.36	40.29	5.76 %

Note: Bold numbers indicate better results.

hyperparameters were automatically optimised through the "fitrgp" function.

Model predictive performance evaluation metrics included mean squared error (MSE), root mean squared error (RMSE), mean absolute error (MAE), and mean absolute percentage error (MAPE). As shown in Table 6, the FNN model had an MSE of 1987.86, RMSE of 44.58, MAE of 33.33, and MAPE of 4.68 %. The GPR model had an MSE of 2966.95, RMSE of 54.36, MAE of 40.29, and MAPE of 5.76 %.

Considering all indicators comprehensively, the FNN model's error metrics on the test dataset were all lower than those of the GPR model. Therefore, the experimental results indicate that under this study's experimental conditions and dataset, the FNN model outperformed the GPR model in the task of cow weight prediction.

3.3.1. Non-pregnant state weight prediction result analysis

To evaluate the accuracy and performance of FNN and GPR models in cow weight prediction, this study conducted model testing on non-pregnant cow. Taking dairy cow numbered cid603 as an example, according to records during the data collection period, this cow did not enter pregnant and had no delivery-related data. This dataset reflects the natural weight changes of dairy cows under normal growth and feeding conditions.

Fig. 13 presents a comparison of the weekly actual weight (green curve, based on measurements obtained from the electronic weighing platform), FNN model predictions (blue curve), GPR model predictions (cyan curve), and farm estimates (magenta curve, calculated based on body length and heart girth measurements) for three regions of dairy cow cid603: dorsal, hips, and side region. The training phase (weeks 1 through 9) shows that the prediction results of both FNN and GPR models are close to the actual weight curve, reflecting the models' ability to fit the training data. The testing phase (weeks 10 through 16, highlighted with a light purple background) is used to evaluate the model's generalisation ability.

Results show that during the testing phase, FNN and GPR models had the most minor prediction errors in the dorsal region, and their prediction stability was superior to that of the hips and side regions. It is particularly noteworthy that the GPR model's prediction results in the dorsal region more closely matched the actual weight trend. Although the GPR model performed slightly worse than the FNN model on the overall training dataset (as mentioned in the previous section), it demonstrated superior generalisation ability on the test dataset for non-pregnant cow. In contrast, farm estimates in all three regions significantly deviated from actual weights, especially during the testing phase, where their curves showed more significant fluctuations and higher errors, highlighting the limitations of traditional body measurement methods regarding accuracy.

The larger prediction errors in the hips region may be related to hips deformation caused by cow movement. Errors in the side region might be affected by the depth camera capturing images in backlit environments; However, the environment was improved before the experiment, and this factor's interference could not be eliminated.

To comprehensively evaluate the predictive performance of the models, this study used four indicators MSE, RMSE, MAE, and MAPE to quantitatively analyse the test phase results of cid603 dairy cow weight prediction. Fig. 14 shows that the GPR model performed best in all regions, especially in the dorsal region, where its MSE, RMSE, MAE, and MAPE values were 3.8e + 02, 19.37, 11.5, and 1.82 %, respectively. The various indicators of the FNN model were slightly higher than those of

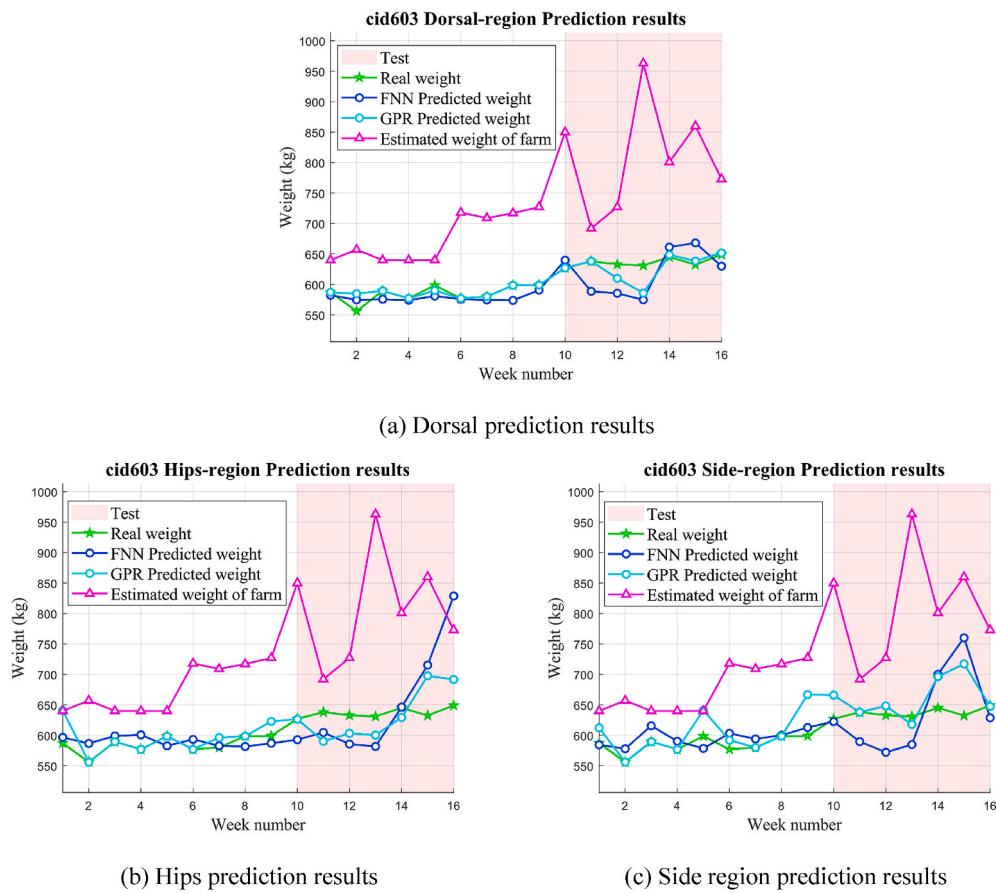


Fig. 13. Weight prediction results for cow ID “cid603”.

the GPR model but still significantly lower than farm estimates. These results further confirm that the farm estimation method (chest girth measurement using a measuring tape) has lower accuracy due to subjective measurements of measuring personnel and limitations of traditional tape measurement methods, mainly showing significant inadequacies in capturing long-term trends of weight fluctuations.

3.3.2. Postpartum state weight prediction result analysis

To verify whether the model could still accurately capture weight trends after cow calving, this study analysed dairy cow numbered cid700 as an example. According to records during the data collection period, this cow was in the pregnant period before (including) week 9 and entered the postpartum stages from week 10 onward.

Fig. 15 presents a comparison of the weekly actual weight (green curve), FNN model predictions (blue curve), GPR model predictions (cyan curve), and farm estimates (magenta curve, calculated based on body length and heart girth measurements) for three regions of dairy cow cid700: dorsal, hips, and side region. The model training phase (weeks 1 through 9) covered the pregnant period. In contrast, the testing phase (weeks 10 through 16, highlighted with a light purple background) corresponded to the postpartum stages, thereby evaluating the models’ generalisation ability when significant changes occur in the physiological state of cow.

The actual weight curve shows a significant decrease after delivery (from week 10), followed by gradual stabilisation, with weight values generally ranging between 650 and 850 kg. This change reflects the weight recovery of dairy cow after delivery and normal fluctuations under feeding conditions. During the training phase, the GPR model’s prediction results for all body regions were consistent with the actual weight trend. The FNN model performed slightly worse in the hips and side regions, failing to accurately capture the relationship between the

features of these regions and actual weight. However, its predictions in the dorsal region were consistent with the actual weight trend.

However, during the testing phase (postpartum stages), both FNN and GPR models showed larger deviations in their prediction results for the hips and side regions. In addition to body part deformations caused by cow movement and backlight issues with the depth camera, these deviations may also stem from changes in cow behaviour patterns after delivery, affecting image features’ stability. In comparison, farm estimates showed more significant deviations during the postpartum stages, with greater curve fluctuations and obvious differences from actual weights, highlighting the limitations of traditional methods in estimating weight under dramatic changes in cow physiological states.

To quantitatively evaluate the models’ performance in predicting dairy cow cid700’s weight during the testing phase (postpartum stages), this study used four indicators MSE, RMSE, MAE, and MAPE for analysis. As shown in Fig. 16, the GPR model performed best in the dorsal region, with MSE, RMSE, MAE, and MAPE values of $5e+02$, 22.35, 18.0, and 2.74 %, respectively. In the hips region, the GPR model’s prediction error was slightly higher than that of the FNN model but still lower than farm estimates. However, in the side region, the prediction errors of both FNN and GPR models were higher than farm estimates, which may reflect the impact of postpartum stages cow behaviour changes and environmental factors on model predictions. Farm estimates (calculated based on body length and heart girth measurements) performed worse in both dorsal and hips regions, with prediction capabilities proving particularly inadequate during phases of intense physiological changes between pregnant and postpartum.

Combining the above results, using relatively stable dorsal region feature data and modelling through the GPR model can effectively capture the nonlinear relationship between feature data and actual weight. Even when cow experience significant weight changes after

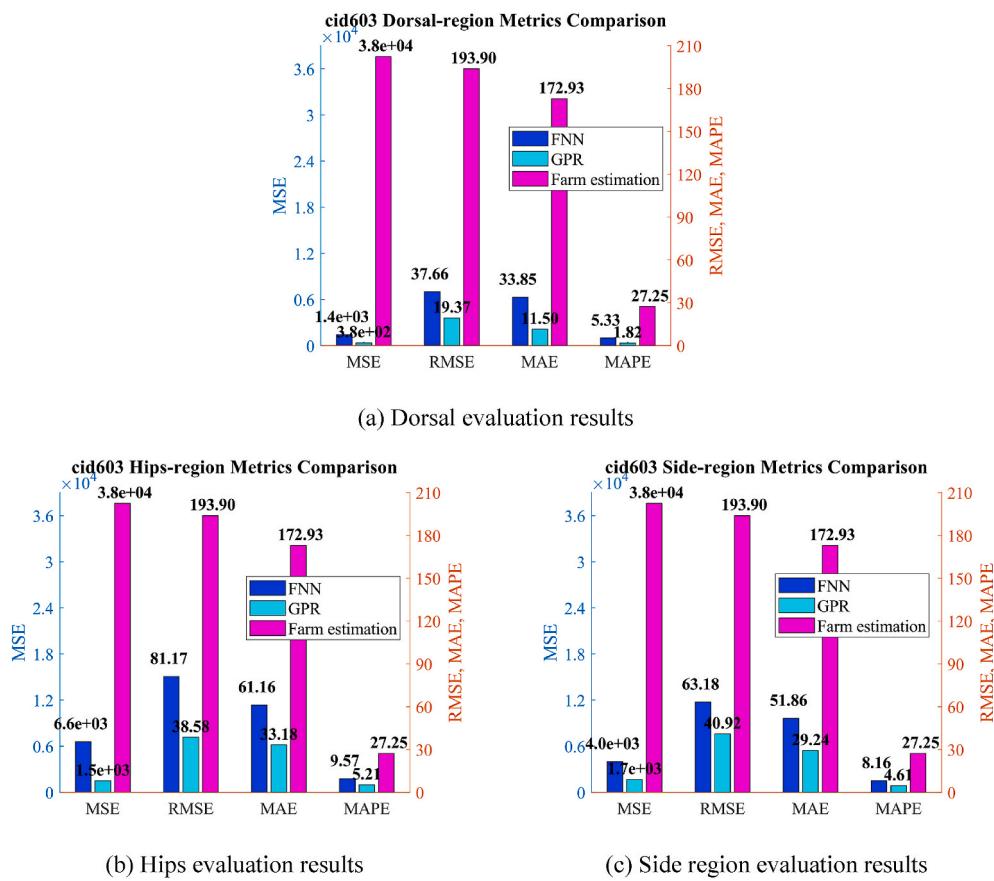


Fig. 14. Weight prediction testing phase results and model evaluation for cow ID “cid603”.

delivery, the GPR model demonstrates good generalisation ability. Although the prediction results for the dorsal region performed excellently in the postpartum stages and the previously discussed pregnant and non-pregnant states, data from multiple angles still needs to be comprehensively analysed to analyse and predict cow weight in a non-contact manner more accurately. Therefore, future research should further optimise and analyse features from various regions to enhance the model’s adaptability and generalisation ability in more complex environments.

3.4. Discussion and research limitations

3.4.1. Methodological comparison with existing research

To objectively evaluate the contributions and limitations of this study, we systematically compare our methodological framework and feature selection strategies with existing similar research. As shown in Table 7, Song et al. (2018) employed Microsoft Kinect v2 sensing equipment, focusing solely on single-view capture from the Hips region, and followed traditional manual selection mechanisms to screen three basic morphological features. This simplified feature extraction approach significantly limits the model’s capability to represent complex body shape variations. In contrast, while Martins et al. (2020) utilised 27 features, they relied on traditional manual feature screening and expert selection practices. Our study integrates ResNet 50-based automated image quality filtering mechanisms, systematic data processing workflows, and morphological feature extraction from three viewing angles—dorsal, Hips, and side—effectively filtering cattle movement noise and environmental interference factors. Dohmen et al. (2021) employed convolutional neural networks to achieve automated feature extraction, breaking through the limitations of traditional manual selection; however, this method lacks representation of animal

morphological features. Relative to the aforementioned research limitations, our study captures dorsal, Hips, and side perspectives through systematic data preprocessing workflows (Fig. 5), establishing ResNet50-based automated image screening mechanisms that effectively filter invalid images caused by noise interference from cattle movement and environmental impacts.

3.4.2. Research limitations

This study identifies four limitations discovered during experimentation that warrant improvement in future research. First, the study employed only five Holstein dairy cows as samples. Although 16 weeks of observational data were collected, encompassing different physiological stages during pregnant and postpartum periods, the sample size remains limited and restricted to a single breed, failing to represent the diversity of dairy cow populations adequately and affecting system generalizability. Second, depth-sensing technology exhibits significant susceptibility to environmental factors. Despite the study’s efforts to reduce environmental light interference through camera installation under eaves and extended support poles, the distinctive black-and-white markings characteristic of Holstein dairy cows readily cause dark areas to absorb infrared light, resulting in depth information loss that requires interpolation processing. This technological limitation may demonstrate varying degrees of impact under different farm environments and climatic conditions. Therefore, future research should consider further incorporating visible light images through depth estimation (Bochkovskiy et al., 2024; Yang et al., 2024) to ensure depth information integrity. Third, the study reveals that prediction accuracy for hip and side regions does not match that of dorsal regions, primarily constrained by shooting angles, body deformation caused by cow movement, and side image interference from backlighting environments—disruptive factors that prove difficult to eliminate completely. Finally, although the

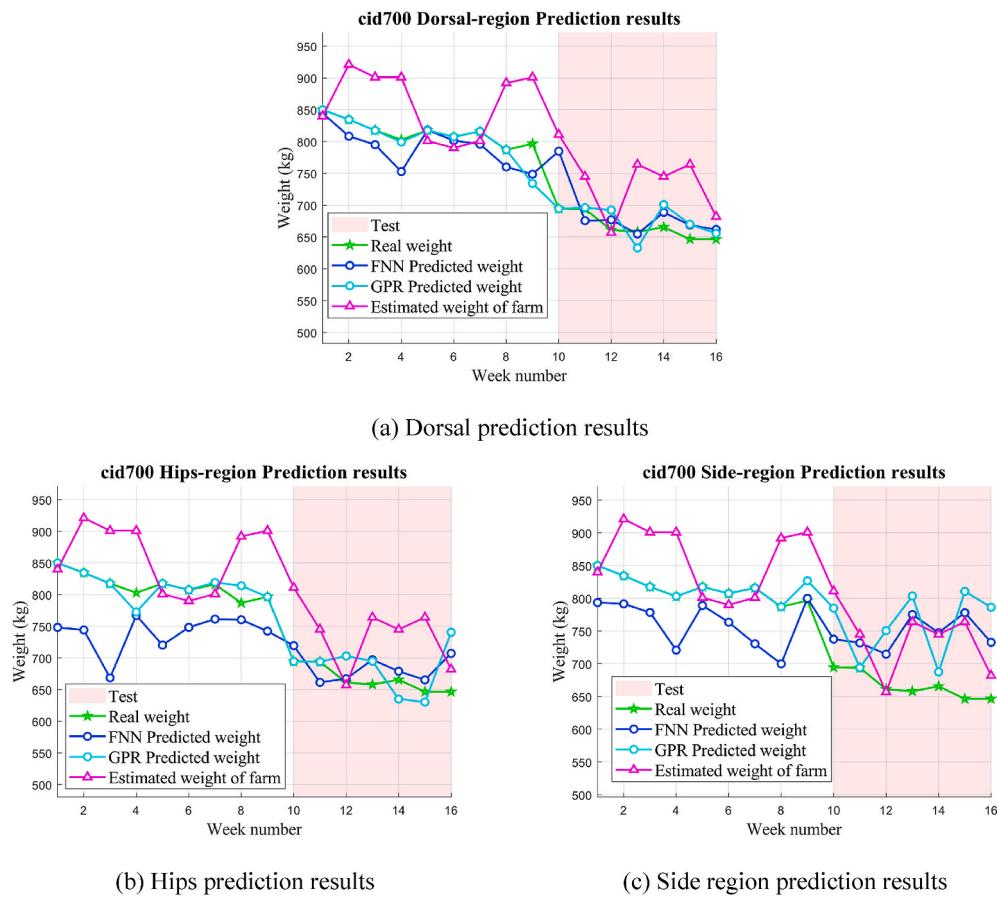


Fig. 15. Weight prediction results for cow ID “cid700” (Postpartum period: Weeks 10–16).

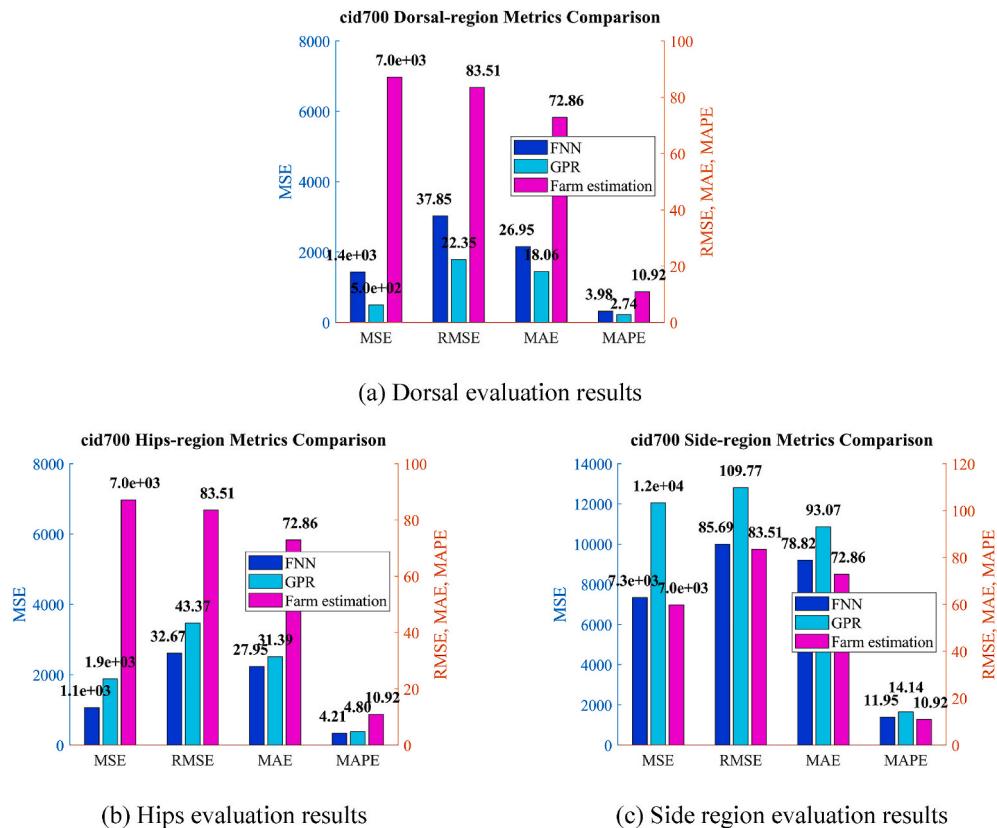


Fig. 16. Weight prediction testing phase results and model evaluation for cow ID “cid700” (Postpartum period: Weeks 10–16).

Table 7
Comparison of feature selection methods.

Study	Sensing Equipment	Capture Region	Feature Selection Method
Song et al. (2018)	Microsoft Kinect v2	Hips	Manual selection of 3 morphological features
Martins et al. (2020)	Microsoft Kinect v2	Dorsal, Side	Manual selection of 27 features
Dohmen et al. (2021)	Microsoft HD-3000 LifeCam	Dorsal, Side	CNN automated feature extraction
Our method	Intel RealSense D455	Dorsal, Hips, Side	Systematic processing + automated feature extraction

system maintains acceptable prediction accuracy during the postpartum period, behavioural pattern changes in postpartum cows (such as increased head movement and tail swaying frequency) significantly impact image feature stability. Moreover, the study encompasses only non-pregnant and postpartum periods, leaving the prediction capability for other critical physiological stages incompletely validated. Future research should consider expanding sample size, implementing systematic ensemble learning (Ruchay et al., 2022b), integrating visible light images for depth estimation to enhance depth information completeness, optimising multi-view image integration techniques, and exploring incorporating additional physiological parameters to improve prediction model robustness.

4. Conclusions

This study developed a non-contact weight prediction system for Holstein dairy cows using Intel RealSense D455 depth cameras, focusing on non-pregnant and postpartum stages. The Gaussian Process Regression (GPR) model demonstrated superior performance when applied to dorsal region features, achieving RMSE values of 19.37 kg (MAPE: 1.82 %) for non-pregnant cows and 22.35 kg (MAPE: 2.74 %) for postpartum cows. These results significantly outperformed both Feedforward Neural Networks and traditional farm estimation methods based on body measurements. The systematic data processing workflow, incorporating ResNet 50-based image classification and morphological feature extraction, effectively filtered environmental noise and movement artifacts. While dorsal region predictions showed excellent accuracy and generalisation capability across different physiological states, prediction performance for hip and side regions remained limited by imaging constraints and cow behavioural factors. Future research should focus on optimising multi-view fusion techniques, incorporating additional physiological parameters, and expanding validation across diverse cattle breeds and farm environments to enhance system robustness and practical applicability.

CRediT authorship contribution statement

Hsin-I Chiang: Writing – review & editing, Supervision, Resources, Conceptualization. **Jia-Ming Zhou:** Writing – review & editing, Writing – original draft, Visualization, Software, Methodology, Investigation, Formal analysis, Data curation. **Wen-Lin Chu:** Writing – review & editing, Supervision, Project administration, Methodology, Investigation, Funding acquisition, Conceptualization.

Data availability statement

The code and preprocessed data provided in this paper will be made publicly available at <https://github.com/wenlinchu/Cows-Weight-Prediction.git> upon publication.

Disclosure on the use of AI in scientific writing

No generative AI or AI-assisted technologies were used during the

preparation of this work.

Funding

This work was supported by the interdisciplinary, integrated research project of The ENABLE Center at National Chung-Hsing University (Project title: “Development of an Intelligent Dairy Cow Management System Based on Objective Body Condition Assessment and Automated Image Recognition”).

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.

Acknowledgements

This work was supported by the interdisciplinary, integrated research project of The ENABLE Center at National Chung-Hsing University. This research was also supported (in part) by NSTC 113-2634-F-005-002 - project Smart Sustainable New Agriculture Research Center (SMARTer). The authors appreciate the funding organizations for their financial support. The authors would also like to thank the helpful comments and suggestions provided by all the researchers cited in this article and the anonymous reviewers. We extend our gratitude to the staff at the practice farm of the Department of Animal Science at National Chung-Hsing University for their assistance in animal management throughout the experimental period.

References

- Bochkovskiy, A., Delaunoy, A., Germain, H., Santos, M., Zhou, Y., Richter, S. R., & Koltun, V. (2024). Depth pro: Sharp monocular metric depth in less than a second. *ArXiv*, abs/2410.02073.
- Carfagni, M., Furferi, R., Governi, L., Servi, M., Uccheddu, F., & Volpe, Y. (2017). On the performance of the Intel SR300 depth camera: Metrological and critical characterization. *IEEE Sensors Journal*, 17(14), 4508–4519. <https://doi.org/10.1109/JSEN.2017.2703829>
- Carroll, C. L., & Huntington, P. J. (1988). Body condition scoring and weight estimation of horses. *Equine Veterinary Journal*, 20(1), 41–45. <https://doi.org/10.1111/j.2042-3306.1988.tb01451.x>
- Carvalho, P. D., Souza, A. H., Amundson, M. C., Hackbart, K. S., Fuenzalida, M. J., Herlihy, M. M., Ayres, H., Dresch, A. R., Vieira, L. M., Guenther, J. N., Grummer, R. R., Fricke, P. M., Shaver, R. D., & Wiltbank, M. C. (2014). Relationships between fertility and postpartum changes in body condition and body weight in lactating dairy cows. *Journal of Dairy Science*, 97(6), 3666–3683. <https://doi.org/10.3168/jds.2013-7809>
- D'Errico, J. (2025). *Inpaint_nans*. MATLAB central file exchange. https://www.mathworks.com/matlabcentral/fileexchange/4551-inpaint_nans.
- Dohmen, R., Catal, C., & Liu, Q. (2021). Image-based body mass prediction of heifers using deep neural networks. *Biosystems Engineering*, 204, 283–293. <https://doi.org/10.1016/j.biosystemseng.2021.02.001>
- Enevoldsen, C., & Kristensen, T. (1997). Estimation of body weight from body size measurements and body condition scores in dairy cows. *Journal of Dairy Science*, 80 (9), 1889–1995. [https://doi.org/10.3168/jds.S0022-0302\(97\)76142-3](https://doi.org/10.3168/jds.S0022-0302(97)76142-3)
- Erickson, P. S., Anderson, J. L., Kalscheur, K. F., Lascano, G. J., Akins, M. S., & Heinrichs, A. J. (2020). Symposium review: Strategies to improve the efficiency and profitability of heifer raising. *Journal of Dairy Science*, 103(6), 5700–5708. <https://doi.org/10.3168/jds.2019-17419>
- Fu, Q., Shen, W., Wei, X., Zheng, P., Xin, H., & Zhao, C. (2019). Prediction of the diet nutrients digestibility of dairy cows using Gaussian process regression. *Information Processing in Agriculture*, 6(3), 396–406. <https://doi.org/10.1016/j.inpa.2018.11.005>
- Ge, Z., Liu, S., Wang, F., Li, Z., & Sun, J. (2021). YOLO: Exceeding YOLO series in 2021. *ArXiv*, abs/2107.08430 <https://doi.org/10.48550/arXiv.2107.08430>
- Hatakeyama, K., Okubo, Y., Nakagome, T., Makino, M., Takashima, H., Akutsu, T., Sawamoto, T., Nagase, M., Noguchi, T., & Kawahito, S. (2023). A hybrid ToF image sensor for long-range 3D depth measurement under high ambient light conditions. *IEEE Journal of Solid-State Circuits*, 58(4), 983–992. <https://doi.org/10.1109/JSSC.2023.3238031>
- Heinrichs, A. J., Erb, H. N., Rogers, G. W., Cooper, J. B., & Jones, C. M. (2007). Variability in Holstein heifer heart-girth measurements and comparison of prediction equations for live weight. *Preventive Veterinary Medicine*, 78(3), 333–338. <https://doi.org/10.1016/j.prevetmed.2006.11.002>

- Heinrichs, A. J., & Hargrove, G. L. (1994). Standards of weight and height for ayrshire, brown Swiss, and milking shorthorn heifers. *Journal of Dairy Science*, 77(6), 1676–1681. [https://doi.org/10.3168/jds.S0022-0302\(94\)77109-5](https://doi.org/10.3168/jds.S0022-0302(94)77109-5)
- Heinrichs, A. J., Heinrichs, B. S., Jones, C. M., Erickson, P. S., Kalscheur, K. F., Nennich, T. D., Heins, B. J., & Cardoso, F. C. (2017). Short communication: Verifying Holstein heifer heart girth to body weight prediction equations. *Journal of Dairy Science*, 100(10), 8451–8454. <https://doi.org/10.3168/jds.2016-12496>
- Kuzuhara, Y., Kawamura, K., Yoshitoshi, R., Tamaki, T., Sugai, S., Ikegami, M., Kurokawa, Y., Obitsu, T., Okita, M., Sugino, T., & Yasuda, T. (2015). 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*, 111, 186–193. <https://doi.org/10.1016/j.compag.2014.12.020>
- Le Cozler, Y., Allain, C., Caillot, A., Delouard, J. M., Delattre, L., Luginbuhl, T., & Faverdin, P. (2019). High-precision scanning system for complete 3D cow body shape imaging and analysis of morphological traits. *Computers and Electronics in Agriculture*, 157, 447–453. <https://doi.org/10.1016/j.compag.2019.01.019>
- Li, Y., Peng, J., Zhang, Y., & Xiong, Z. (2023). Self-distilled depth from single-shot structured light with intensity reconstruction. *IEEE Transactions on Computational Imaging*, 9, 678–691. <https://doi.org/10.1109/TCI.2023.3289596>
- Lu, J., Guo, H., Du, A., Su, Y., Ruchay, A., Marinello, F., & Pezzuolo, A. (2022). 2-D/3-D fusion-based robust pose normalisation of 3-D livestock from multiple RGB-D cameras. *Biosystems Engineering*, 223, 129–141. <https://doi.org/10.1016/j.biosystemseng.2021.12.013>
- Martins, B. M., Mendes, A. L. C., Silva, L. F., Moreira, T. R., Costa, J. H. C., Rotta, P. P., Chizzotti, M. L., & Marcondes, M. I. (2020). Estimating body weight, body condition score, and type traits in dairy cows using three dimensional cameras and manual body measurements. *Livestock Science*, 236, Article 104054. <https://doi.org/10.1016/j.livsci.2020.104054>
- Martinson, K., Coleman, R., Rendahl, A., Fang, Z., & McCue, M. (2014). Estimation of body weight and development of a body weight score for adult equids using morphometric measurements. *Journal of Animal Science*, 92(5), 2230–2238. <https://doi.org/10.2527/jas.2013-6689>
- Milner, J., & Hewitt, D. (1969). Weight of horses: Improved estimates based on girth and length. *Canadian Veterinary Journal*, 10(12), 314–316. <https://pubmed.ncbi.nlm.nih.gov/5392700/>
- Ojha, V. K., Abraham, A., & Snásel, V. (2017). Metaheuristic design of feedforward neural networks: A review of two decades of research. *Engineering Applications of Artificial Intelligence*, 60, 97–116. <https://doi.org/10.1016/j.engappai.2017.01.013>
- Preethi, A. L., Tarafdar, A., Ahmad, S. F., Panda, S., Tamilarasan, K., Ruchay, A., & Gaur, G. K. (2023). Weight prediction of landly pigs from morphometric traits in different age classes using ANN and non-linear regression models. *Agriculture*, 13(2).
- Rico, J. E., Bandaru, V. V. R., Dorskind, J. M., Haughey, N. J., & McFadden, J. W. (2015). Plasma ceramides are elevated in overweight Holstein dairy cows experiencing greater lipolysis and insulin resistance during the transition from late pregnant to early lactation. *Journal of Dairy Science*, 98(11), 7757–7770. <https://doi.org/10.3168/jds.2015-9519>
- Rodríguez Alvarez, 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. *Computers and Electronics in Agriculture*, 155, 12–22. <https://doi.org/10.1016/j.compag.2018.09.039>
- Rotondo, V., Osborne, V. R., Paibomesai, M., Wood, K. M., & Jantzi, S. (2021). PSX-A-21 late-breaking: Predicting live weight using linear body measurements in growing dairy calves. *Journal of Animal Science*, 99(Supplement_3), 371–372. <https://doi.org/10.1093/jas/skab235.681>
- Ruchay, A., Gritsenko, S., Ermolova, E., Bochkarev, A., Ermolov, S., Guo, H., & Pezzuolo, A. (2022). A comparative study of machine learning methods for predicting live weight of Duroc, landrace, and yorkshire pigs. *Animals*, 12(9).
- Ruchay, A., Kober, V., Dorofeev, K., Kolpakov, V., Dzhulamanov, K., Kalschikov, V., & Guo, H. (2022b). Comparative analysis of machine learning algorithms for predicting live weight of Hereford cows. *Computers and Electronics in Agriculture*, 195, Article 106837. <https://doi.org/10.1016/j.compag.2022.106837>
- Ruchay, A., Kober, V., Dorofeev, K., Kolpakov, V., Gladkov, A., & Guo, H. (2022c). Live weight prediction of cattle based on deep regression of RGB-D images. *Agriculture*, 12 (11).
- Sharpe, K. T., & Heins, B. J. (2023). Evaluation of a forefront weight scale from an automated calf milk feeder for holstein and crossbred dairy and dairy-beef calves. *Animals*, 13(11), Article 1752. <https://doi.org/10.3390/ani13111752>
- Shen, W., Li, G., Wei, X., Fu, Q., Zhang, Y., Qu, T., Chen, C., & Wang, R. (2022). Assessment of dairy cow feed intake based on BP neural network with polynomial decay learning rate. *Information Processing in Agriculture*, 9(2), 266–275. <https://doi.org/10.1016/j.inpa.2021.04.008>
- Silva, F. G., Carreira, E., Ramalho, J. M., Correia, T., Meira, M., Conceição, C., Silva, S. R., Pereira, A. M. F., & Cerqueira, J. L. (2024). Predicting body weight in pre-weaned holstein-friesian calves using morphometric measurements. *Animals*, 14 (14), Article 2129. <https://doi.org/10.3390/ani14142129>
- Song, X., Bokkers, E. A. M., van der Tol, P. P. J., Groot Koerkamp, P. W. G., & van Mourik, S. (2018). Automated body weight prediction of dairy cows using 3-dimensional vision. *Journal of Dairy Science*, 101(5), 4448–4459. <https://doi.org/10.3168/jds.2017-13094>
- Wagner, E. L., & Tyler, P. J. (2011). A comparison of weight estimation methods in adult horses. *Journal of Equine Veterinary Science*, 31(12), 706–710. <https://doi.org/10.1016/j.jevs.2011.05.002>
- Wang, H., Che, Z., Yang, Y., Wang, M., Xu, Z., Qiao, X., Qi, M., Feng, F., & Tang, J. (2024). RDGC-GAN: RGB-depth fusion CycleGAN for indoor depth completion. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 46(11), 7088–7101. <https://doi.org/10.1109/TPAMI.2024.3388004>
- Wang, Z., Shadpour, S., Chan, E., Rotondo, V., Wood, K. M., & Tulpan, D. (2021). ASAS-NANP symposium: Applications of machine learning for livestock body weight prediction from digital images. *Journal of Animal Science*, 99(2), skab022. <https://doi.org/10.1093/jas/skab022>
- Wangchuk, K., Wangdi, J., & Mindu, M. (2018). Comparison and reliability of techniques to estimate live cattle body weight. *Journal of Applied Animal Research*, 46(1), 349–352. <https://doi.org/10.1080/09712119.2017.1302876>
- Wong, T. T., & Yeh, P. Y. (2020). Reliable accuracy estimates from k-Fold cross validation. *IEEE Transactions on Knowledge and Data Engineering*, 32(8), 1586–1594. <https://doi.org/10.1109/TKDE.2019.2912815>
- Yan, T., Mayne, C. S., Patterson, D. C., & Agnew, R. E. (2009). Prediction of body weight and empty body composition using body size measurements in lactating dairy cows. *Livestock Science*, 124(1), 233–241. <https://doi.org/10.1016/j.livsci.2009.02.003>
- Yang, L., Kang, B., Huang, Z., Zhao, Z., Xu, X., Feng, J., & Zhao, H. (2024). Depth anything V2. *ArXiv abs/2406.09414*.
- Zhao, K., Zhang, M., Shen, W., Liu, X., Ji, J., Dai, B., & Zhang, R. (2023). Automatic body condition scoring for dairy cows based on efficient net and convex hull features of point clouds. *Computers and Electronics in Agriculture*, 205, Article 107588. <https://doi.org/10.1016/j.compag.2022.107588>