



CNN-based weight estimation from point clouds obtained from walking breed sows

Kiyoun Kwon^a, Jun Hwan Park^b, Ahram Park^a, Sangwoo Kim^a, Nojun Lee^a, Duhwan Mun^{b,*}

^a School of Industrial Engineering, Kumoh National Institute of Technology, 61 Daehak-ro, Gumi, Gyeongbuk, 39177, South Korea

^b School of Mechanical Engineering, Korea University, 145 Anam-ro, Seongbuk-gu, Seoul, 02841, South Korea

ARTICLE INFO

Keywords:

CNN
Breed sows
Point cloud
Regularized image
Two-dimensional mesh
Weight estimation

ABSTRACT

In the livestock industry, reducing the cost of raising animals while maintaining quality is of utmost importance. Growth management is crucial for livestock quality control, with weight being a primary focus. This study proposes a method for estimating the weight of a breed sow from point clouds using a convolutional neural network (CNN). The proposed method consists of two stages: first, 2D images are generated from point clouds of a walking breed sow; next, these images are used as input in a CNN model to estimate the weight. The point clouds are divided into two halves along a longitudinal plane. Then, distances are calculated between each element of the 128×64 rectangular mesh created on the aforementioned plane and the corresponding points in the divided point cloud. A 2D image is generated by varying the grayscale intensity according to the calculated distances. Three types of CNNs, namely, VGG16, DenseNet121, and EfficientNet B0, were employed to estimate the weight using the generated 2D images as input. The CNNs were trained using approximately 150,000 images of 71 pigs. To verify the effectiveness of the method, the trained networks were used to conduct weight estimation experiments on test cases. The results of the experiments demonstrate a high weight estimation accuracy, with an error rate as low as 1.35%.

1. Introduction

With the continuous increase in meat consumption, reducing the cost of raising livestock while maintaining high-quality standards is essential to ensure competitiveness in the livestock industry, particularly in pig farming. Consequently, the establishment of smart farms has gained significant attention within the livestock sector. Effective growth control plays a pivotal role in managing the quality of livestock, with weight being a key parameter [1–3]. Weight serves as a valuable indicator of how animal health and growth are influenced by factors such as feed and farm environments, i.e., it enables the evaluation of pigs' growth rate and feed conversion efficiency [4]. Currently, livestock weights are predominantly measured using scales, a process that demands considerable time and effort and can lead to errors due to the movement of livestock [5].

The method of directly measuring livestock weight using scales or estimating weight by manually measuring related body dimensions with a ruler is widely used [6]. However, this method consumes significant labor and can have negative impacts, such as causing stress to the livestock or injuries to both workers and the animals [7,8]. These

problems have been highlighted in various studies, and different approaches have been analyzed to address these issues [9–11].

To overcome the issues of direct measurement methods, indirect methods are actively being researched. These methods extract physical characteristics of livestock from images and apply mathematical regression analysis, or they use deep learning to automatically learn physical characteristics for weight estimation. These studies used single photos [12–14], depth images [15–21], point clouds [22–24], and mesh models [25,26]. However, these studies often require complex setups, such as multiple cameras, to acquire data [27], which necessitates calibration [28], and measurements must be taken from fixed positions, leading to issues with mobility and increased costs. Moreover, capturing the belly region, which significantly affects the weight of breed sows, is a challenging task due to the movement of animals [29] and the limitations of measurement systems. In some cases, RGB and RGB-D cameras are used, but they often generate unnecessary data that complicates accurate weight estimation. Although RGB-D images provide depth information, they often include noise that degrades performance.

Furthermore, traditional weight estimation methods, such as those using regression equations [15,17,25], tend to rely on selected key body

* Corresponding author.

E-mail address: dhmun@korea.ac.kr (D. Mun).

<https://doi.org/10.1016/j.atech.2024.100611>

Received 9 July 2024; Received in revised form 10 October 2024; Accepted 17 October 2024

Available online 18 October 2024

2772-3755/© 2024 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

characteristics rather than considering the overall body shape of livestock. This limited approach can reduce weight estimation accuracy, and the results can vary depending on the selection of characteristics and how weights are assigned. Recently, deep learning has emerged as an alternative that more comprehensively analyzes the overall body shape of livestock. Studies have employed artificial neural networks (ANN) [20,26] and convolutional neural networks (CNN) [18,24,30–32]. However, many of these studies focus on growing and finishing pigs and sometimes show relatively lower performance.

In this study, we propose a CNN-based method for estimating breed sow weight using partially measured point clouds. The proposed method consists of two stages: first, 2D images are generated from point clouds of a walking breed sow; next, these images are used as input in a CNN model to estimate the weight. The point clouds are divided into two halves along a longitudinal plane. Then, distances are calculated between each element of the 128×64 rectangular mesh created on the aforementioned plane and the corresponding points in the divided point cloud. A 2D image is generated by varying the grayscale intensity according to the calculated distances. Three types of CNNs, namely, VGG16, DenseNet121, and EfficientNet B0, are employed to estimate the weight using the generated 2D images as input. The process is illustrated in Fig. 1.

This study makes the following academic contributions. First, it proposes a CNN-based weight estimation method, showcasing state-of-the-art performance with a mean absolute percentage error (MAPE) of 1.35. Second, unlike previous studies targeting growing and finishing pigs, it focuses on breed sows, including the belly part in weight estimation. Third, it devises a generation process for 2D grayscale images

optimized for weight estimation.

The structure of this paper is as follows: Section 2 reviews the data types and various approaches used for livestock weight estimation. Section 3 explains the data generation methods used in the experiments, the three CNN models (VGG16, DenseNet121, and EfficientNet B0) employed, and the experimental results. Section 4 covers the analysis of the experimental results and discusses potential improvements. Finally, Section 5 provides the conclusion of this study and discusses future study directions.

2. Related works

The live weight of livestock is a crucial indicator for monitoring daily weight gain, nutritional status, health, as well as predicting and managing market shipment weights [9]. Traditionally, livestock weight is measured directly using scales [9–11], or estimated based on empirical relationships between morphological characteristics and weight [33], or by measuring body dimensions with a ruler to estimate the weight [6, 34]. However, direct measurement methods require different scales for animals with varying body shapes, and there is the inconvenience of having to accurately position animals on the scales. Furthermore, these methods demand a significant amount of labor, and during the measurement process, both the livestock and the handlers may experience stress, which poses a risk of injury [8,35]. Due to these issues, along with increasing market demand, the automated management systems has been recommended [36]. Consequently, approaches that minimize contact with livestock, using linear regression algorithms or deep learning algorithms, are being actively researched.

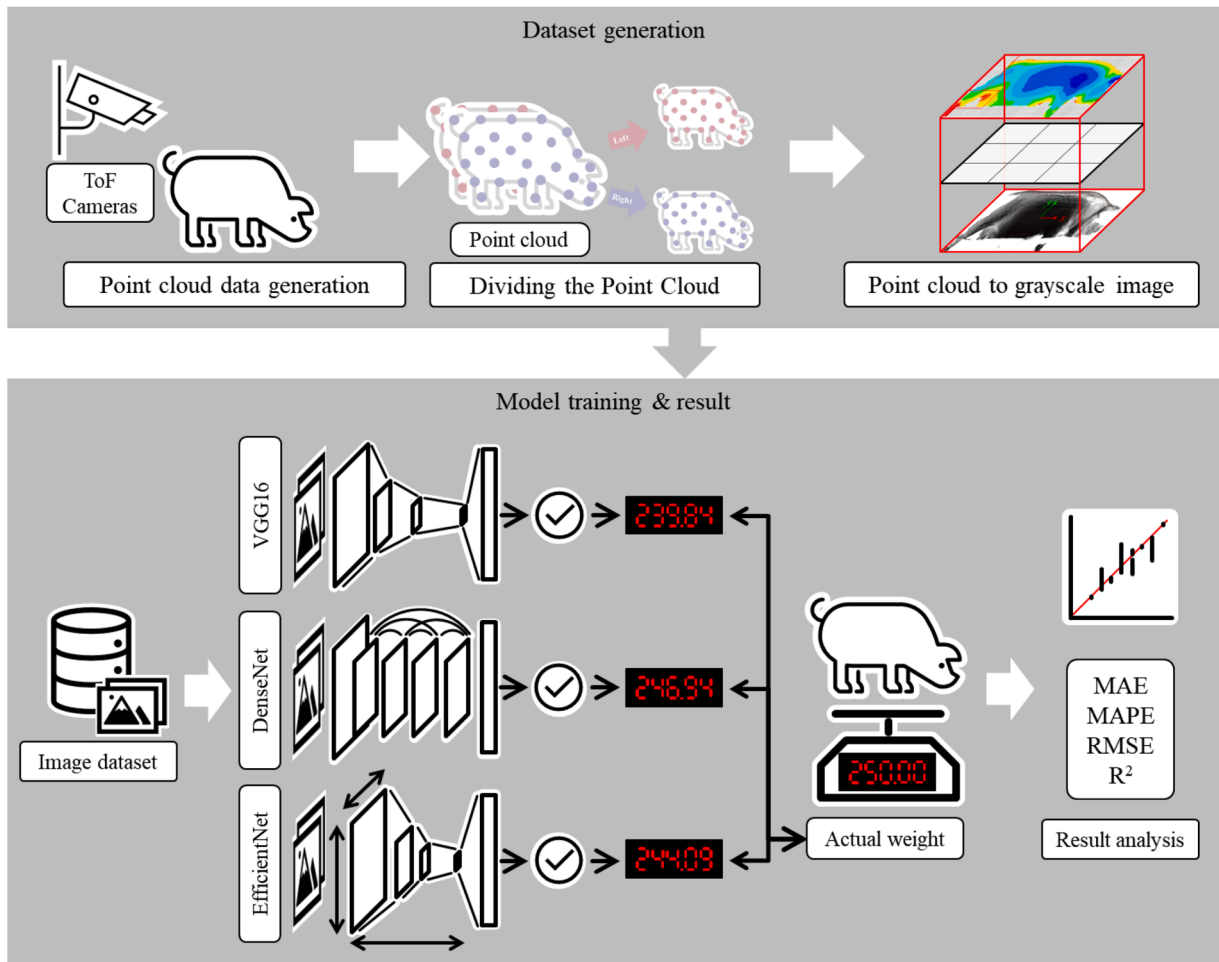


Fig. 1. Overall process of CNN-based breed sow weight estimation.

Previous studies that used linear regression and deep learning algorithms have utilized various types of data for predicting livestock weight, including single photos [12–14], depth images [15–21], point clouds [22–24], and mesh models [25,26]. Point clouds are typically obtained through depth scanners or laser scanners and can be directly employed as weight estimation data or reconstructed into mesh models. For instance, Kashiha et al. [12] developed a method that estimated the weight of pigs in a pen by identifying the areas where pigs are present based on images taken with a top-view camera. Shi et al. [13] created a mobile measurement system to gather data for measuring the main body dimensions (body length, body width, body height, hip width, and hip height) of pigs in a large-scale farm setting. Pezzuolo et al. [15] estimated pig weights by calculating their main dimensions using two low-cost 3D cameras and inputting the data into a weight estimation model. Similarly, Pezzuolo et al. [16] devised a system utilizing a structured light depth camera to monitor cow growth, estimating the cows' body sizes based on the main dimensions calculated from point clouds. In other studies, Conodotta et al. [17] predicted pig weight by generating volumetric information using a single RGB-D camera. Wang et al. [22] proposed an innovative algorithm for measuring body size using a portable system equipped with dual-depth cameras, which extracted segmented point clouds of pigs from the acquired data. Cozler et al. [23] developed a method for analyzing the body shapes of cows by registering multiple acquired point clouds to create a unified point cloud. Nguyen et al. [21] employed a CNN to segment depth images in the process of pig weight estimation. Kwon et al. [25,26] predicted pig weights by constructing mesh models from point clouds obtained through multiple RGB-D cameras. Shuai et al. [19] introduced a method for measuring the body size of freely moving pigs by capturing point clouds from multi-view RGB-D cameras.

Methods for estimating livestock weight without direct measurement can be divided into traditional image-based measurement methods and deep learning algorithm-based approaches. In the image-based methods, livestock is photographed using devices such as RGB cameras, 3D time-of-flight (ToF) cameras [10], or LiDAR [37], and then physical characteristics of the livestock are extracted. The most commonly used features include top view body area, withers height, hip height, body length, hip width, body volume, and chest girth [10]. Various linear regression algorithms are then used to estimate the weight of the livestock. Al Ard Khanji et al. [38] used flank-to-flank distance, length, heart girth, ultrasound backfat measurement, loin depth, and body condition score. They applied a linear regression algorithm using these six body measurements and the pigs' physiological conditions, achieving an R^2 of 0.90. Shi et al. [39] generated pig body contours and reconstructed pig images through image analysis and obtained the back area of pigs. They then determined the weight estimation function according to the regression results using the least squares method and achieved an R^2 of 0.9931. However, methods using linear regression still require manual extraction and selection of physical characteristics [11], and there are limitations since the relationship between input and output variables must be predefined [9]. To address these issues, approaches using deep learning have been researched.

Deep learning is a machine learning method that learns the data representation by itself [19]. Condotta et al. [20] applied regression equations and ANNs using depth images to estimate pig weight. However, the prediction accuracy of these two studies, which focused on growing and finishing pigs, was not notably high. Kwon et al. [26], estimated pig weights using multiple depth cameras. They reconstructed a 3D shape from the acquired point cloud and measured the main dimensions to input into an ANN for weight estimation, achieving a 2% error rate and accurate sow weight prediction. In addition to ANNs, CNNs, which excel at learning image features autonomously, have been widely used for livestock weight estimation. Ruchay et al. [18] used a CNN to predict cow weight based on depth images acquired using an RGB-D camera, achieving a prediction performance of MAPE 8.4. Meckbach et al. [30], He et al. [31], and Liu et al. [32] applied a CNN to

predict weight for growing and finishing pigs using top-view data. However, their prediction performances were relatively low, with MAPE values of 3.9, 6.37, and 2.82, respectively. Liu et al. [40] used a hybrid 3D point cloud denoising approach for accurate pig weight prediction. After using six parameters from the pigs' back as input to the CNN, they achieved a mean absolute error (MAE), MAPE, and root mean square error (RMSE) of 12.45 kg, 5.36%, and 12.91 kg, respectively, proving the effectiveness of voxel downsampling and point cloud techniques. Dohmen et al. [41] used Mask-RCNN for image segmentation and background removal and employed CNN to predict calf weight, achieving R^2 of 0.96 and RMSE of 20 kg from top-view images. Bi et al. [42] explored long-term cattle weight prediction using depth video data and achieved an R^2 of 0.98 and MAPE of 2.03% by combining the Mask-RCNN approach with a linear mixed model. Gjergji et al. [43] analyzed various deep learning architectures, including CNNs, recurrent neural networks (RNN)/CNN networks, and recurrent attention models. As a result, convolutional neural networks showed the best performance in predicting cattle weight, reducing the error of traditional linear regression models by almost half, with an MAE of 23.19 kg.

3. Methodology

3.1. Animals

The data utilized in this study, which proceeded from a previous work [26], were divided into two halves: left and right. A total of 71 sows, with weights between 168 and 313 kg, were included in the experiments. A point cloud acquisition system was installed, as shown in Fig. 2. Four ToF cameras (Azure Kinect, Microsoft) were used as the measuring equipment. The four cameras were fixed to separate camera frames, and guide frames were installed to secure the field of view (FoV) for the cameras. The width of the corridor was 2 m, and the two upper cameras were installed at the height of 2 m to allow people to pass through.

The dataset comprised 1022 measurement data points, as shown in Fig. 3. Among these data points, 1631 containing measurements of both hind and forelegs were selected for this study. Fig. 3(a) presents the measured data for an entire pig, whereas Fig. 3(b) shows the data divided into the left and right sections.

Fig. 4 shows an example of the point clouds used for training. No points were considered in the buttock area or in most parts of the head, as shown in Fig. 4(a). Additionally, parts of the legs are obscured by the frame of the measurement system and thus are not present. Furthermore, as shown in the enlarged insets of Fig. 4(b), a significant amount of noise is also present.

3.2. Dataset construction

The process of image transformation involves creating a regularized grayscale image that takes into account the size and depth information derived from the point cloud. The transformation method is illustrated in Fig. 5. Initially, the point cloud data is input to compute an average plane and cylinder based on which the XY-plane is defined. A reference plane is established, and quadrilateral elements of uniform size are generated on it. The distances between the points and elements are calculated and the resulting values are used to assign grayscale colors.

3.2.1. Axes transformation

The measurements to create the input point cloud were acquired as the pigs walked through a passageway and using an arbitrary coordinate system, which resulted in slightly different directions from tail to head. To achieve a consistent representation in the 2D image, the direction from the head to tail was set as the X-axis, whereas the direction from the legs to the upper body was defined as the Y-axis, with the origin corresponding to the average coordinates of all points. In this setup, the pig's torso was positioned at the center of the image. Plane fitting was

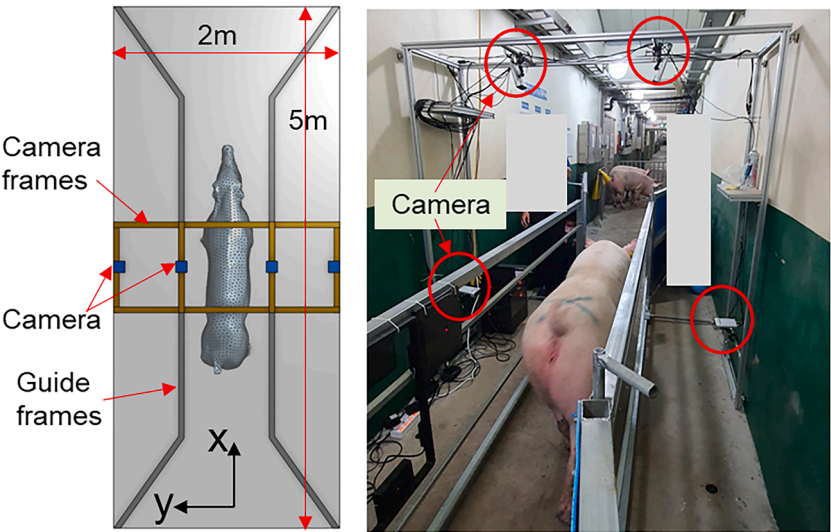


Fig. 2. Point cloud acquisition system configuration.

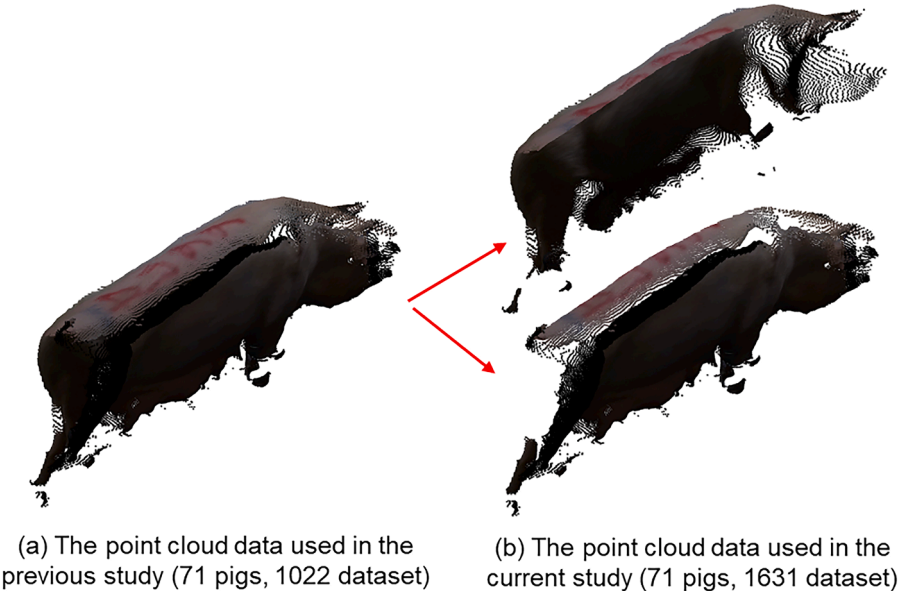


Fig. 3. Point cloud data used for training.

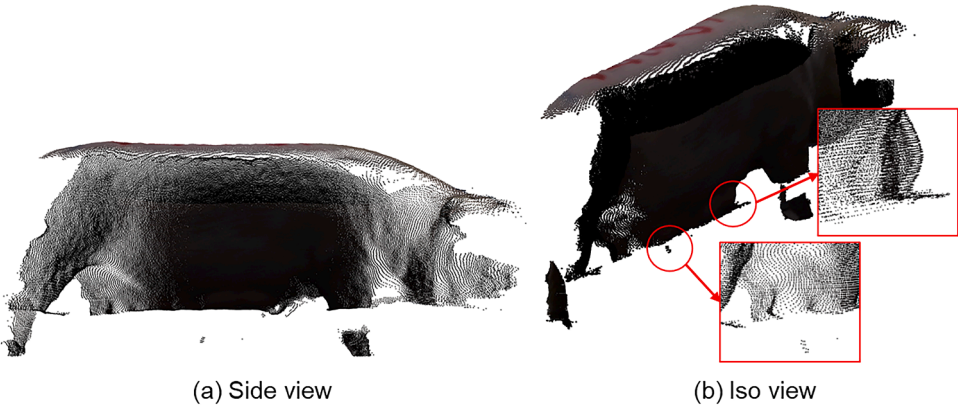


Fig. 4. Characteristics of the point cloud data.

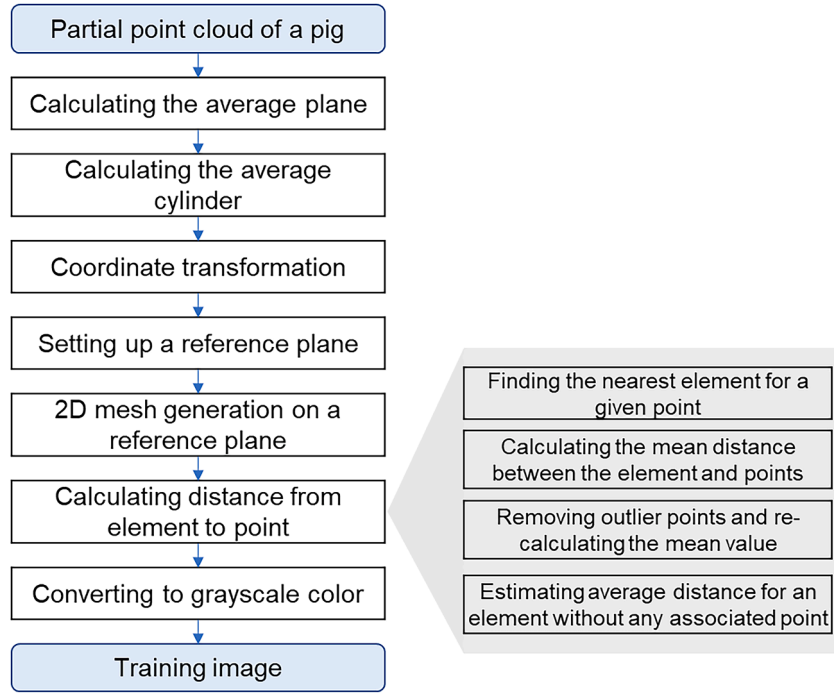


Fig. 5. Transformation of a point cloud into an image.

performed using the entire point cloud, setting the normal direction of the plane as the Z-axis. Cylinder fitting was then conducted to set the direction of the cylinder as the X-axis. Fig. 6 shows the coordinate transformation process. Fig. 6(a) shows the input data, whereas Fig. 6(b) shows the result of plane fitting using all points, where the plane's normal direction aligns with the Z-axis. Fig. 6(c) presents the result of cylinder fitting using all points, with the cylinder's axis direction defining the x-axis. Both plane and cylinder fittings were performed by minimizing the RMSE, expressed as:

$$RMSE = \sqrt{\frac{\sum_{i=1}^N \overline{P_i \hat{P}_i}^2}{N}} \quad (1)$$

where $\overline{P_i \hat{P}_i}$ is the distance from the measurement point (P) to the projected point ($\hat{P}$) on the plane and cylinder. Fig. 6(d) shows the outcome after moving the data to the XY plane.

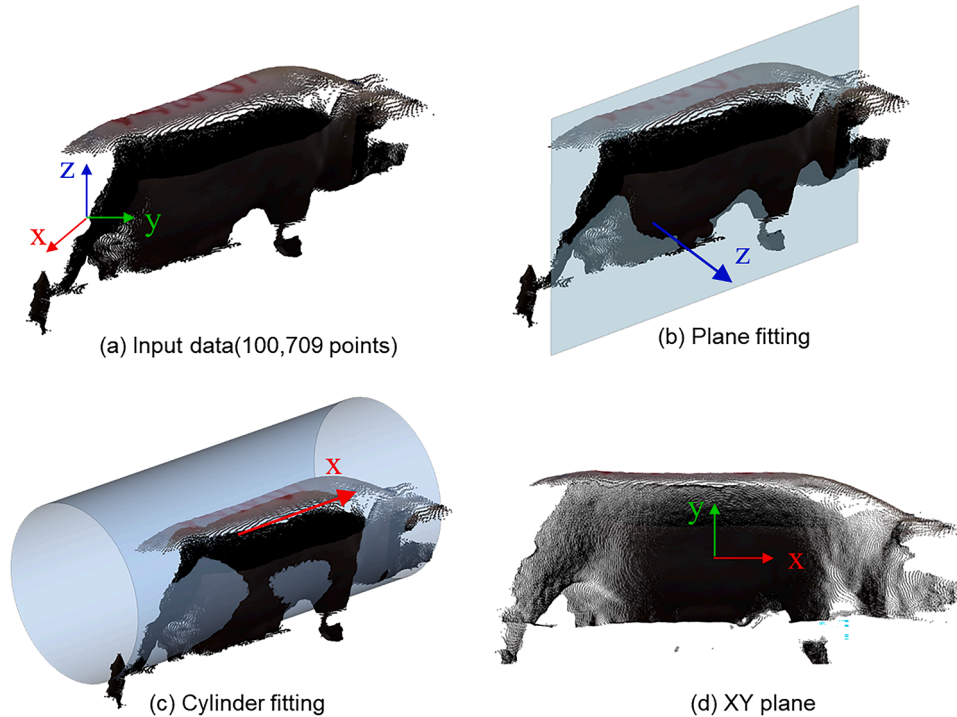


Fig. 6. Transformation of coordinate axes to the XY-plane.

3.2.2. Generation of a regularized image

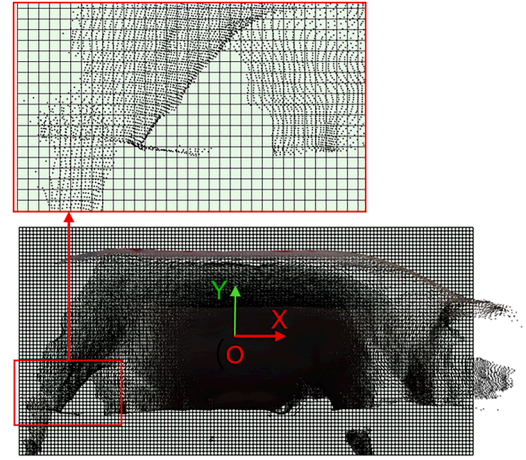
The maximum size of the sow was approximately 2 m in length from tail to head and approximately 1 m in height. However, the legs, tail, and head sections do not contribute significantly to the overall weight, and obtaining accurate data for these parts is relatively challenging compared with the body section. In the present study, the data lacked an adequate number of points for the tail and head sections. The size of the image was set to 1800×900 mm to ensure that the area between the front and hind legs was included. The image used for the training was set to 128×64 pixels with a quadrilateral element size of 14 mm. Consequently, the overall image size was 1792×896 mm, as shown in Fig. 7.

Fig. 8 shows the generation of 128 elements in the x-direction and 64 elements in the y-direction. The element size was 14 mm, and the dimensions were $14 \text{ mm} \times 14 \text{ mm}$. Each square element had a size of 14 mm. In the final transformed image, each quadrilateral element corresponded to one pixel. The elements that encompassed each point across all the points were determined using Eq. (2).

$$\begin{aligned} N_{xi} &= \text{floor}((P_{xi} - X_{Min}) \times N_x / (X_{Max} - X_{Min})) \\ N_{yi} &= \text{floor}((P_{yi} - Y_{Min}) \times N_y / (Y_{Max} - Y_{Min})) \end{aligned} \quad (2)$$

N_{xi} and N_{yi} represent the positions of the x- and y-directional elements at the i th point, respectively. where P_{xi} and P_{yi} denote the x- and y-directional coordinates of the i th point, respectively. Moreover, N_x and N_y indicate the number of elements in the X and Y directions, respectively; N_x is 128 and N_y is 64. Each element stores the encompassed points, and if the count falls below a predetermined threshold, the point is deemed as noise and is subsequently removed. In a single element, approximately ten points were stored; in this study, elements with fewer than three points were excluded. The distance from each point stored in the element to the reference plane is calculated. The reference plane was defined to ensure that points arranged in the z-direction possessed greater values to clearly distinguish between empty regions and segments with measured points in the final training image. The maximum width of each pig is 500 mm. The reference plane was set at a location shifted by 200 mm in the z-direction from the average plane of the partial measurement data (Fig. 9). Given this configuration, the minimum distance (d_{min}) between the reference plane and the point can have a positive value. When half of the data from the opposite side was used, the reference plane was set in the z-direction to yield a positive value.

For each point in an element, the distance to the reference plane was calculated and 30% of the points that deviated from the mean were excluded. Fig. 10 shows an example of distance calculations in two dimensions. A reference plane encompassing all the points was created, and four elements were generated on this reference plane (Fig. 10(a)).



(a) 2D grid generation

Fig. 8. Generating the quadrilateral mesh.

Fig. 10(b) shows the distance from the points within each element to the reference plane. Points that were significantly distant were deemed outliers and excluded before calculating the average distance (Fig. 10(c)). For elements devoid of related points, the average movement distance of the adjacent elements was calculated. Fig. 10(d) shows that the third element does not contain any points, and the average value of +1, derived from the adjacent elements' distance values of +4 and -2, is inserted. When multiple elements are consecutively empty, the value is calculated using the distance ratio of the adjacent elements.

Fig. 11(a) shows the calculated distance values, with a minimum distance of 108.18 mm and a maximum distance of 383.13 mm. The grayscale value can be calculated by dividing the height value by 400, as shown in Eq. (3).

$$\begin{aligned} V_{GC} &= \frac{\text{Height}}{400} \\ \text{if } V_{GC} > 1 \text{ then } V_{GC} &= 1 \\ \text{if } V_{GC} < 0 \text{ then } V_{GC} &= 0 \end{aligned} \quad (3)$$

A value of 400 was set, such that cases distant from the reference plane would assume a value of 1. For exception handling, values less than 0 were set to 0 and those greater than 1 were set to 1. Fig. 11(b) shows the final result converted into a grayscale image, in which the

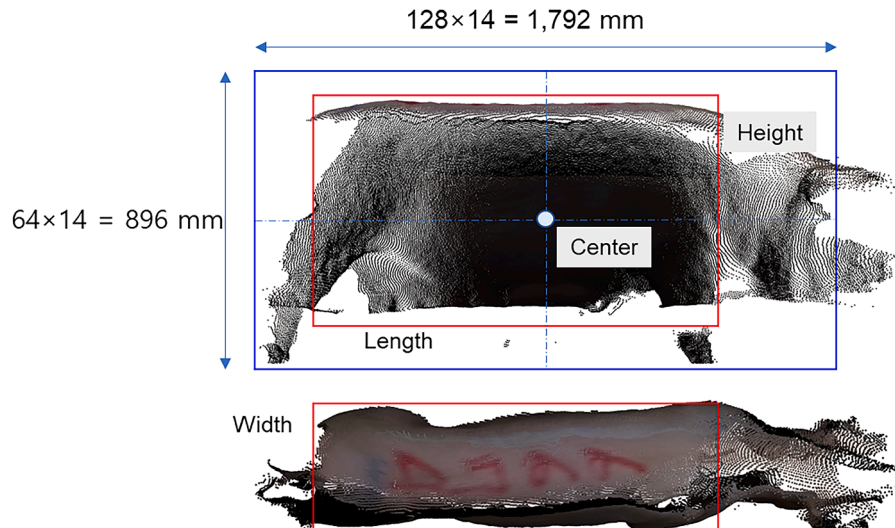


Fig. 7. Determining the image size.

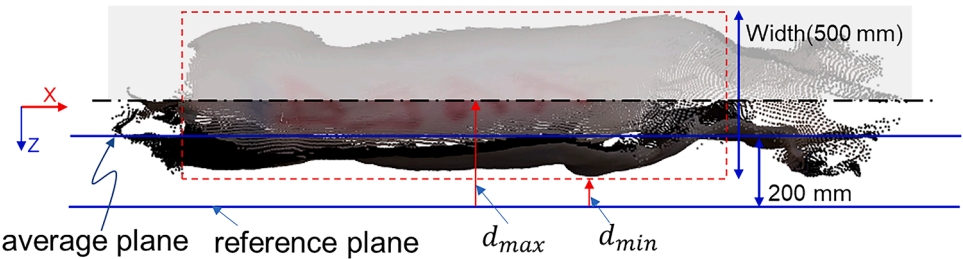


Fig. 9. Setting the reference plane.

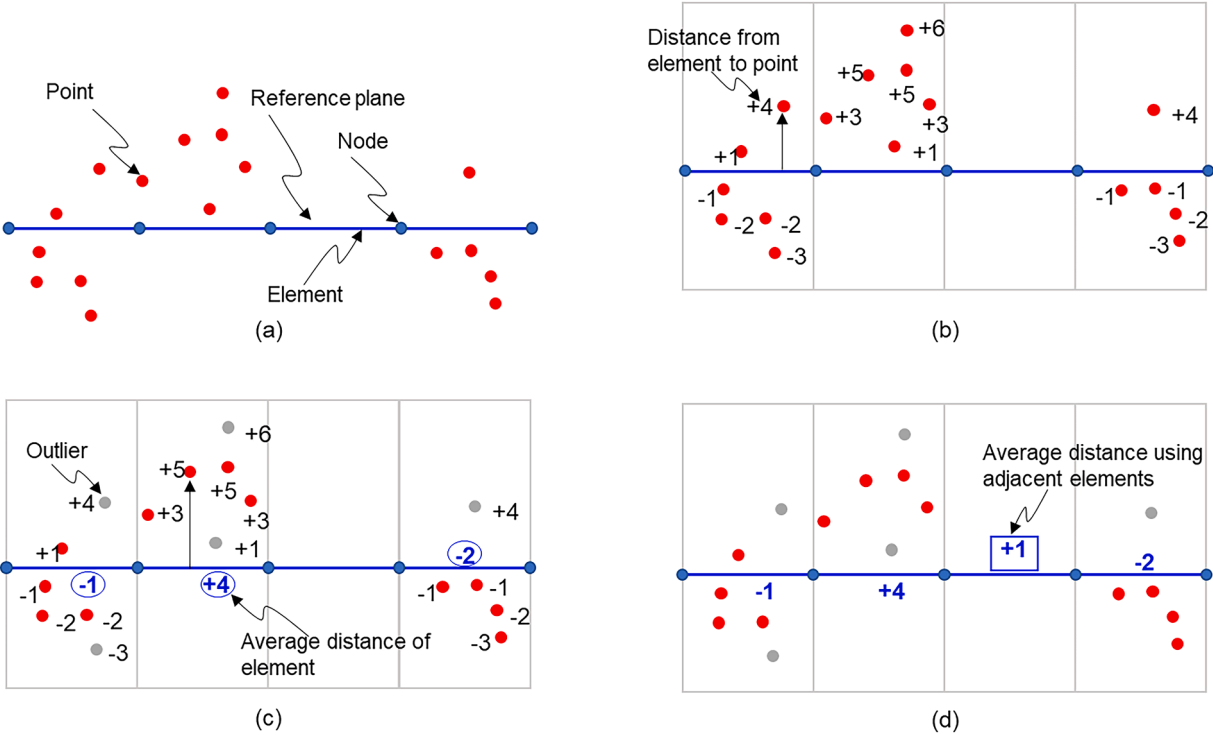


Fig. 10. Calculating distances between points and elements.

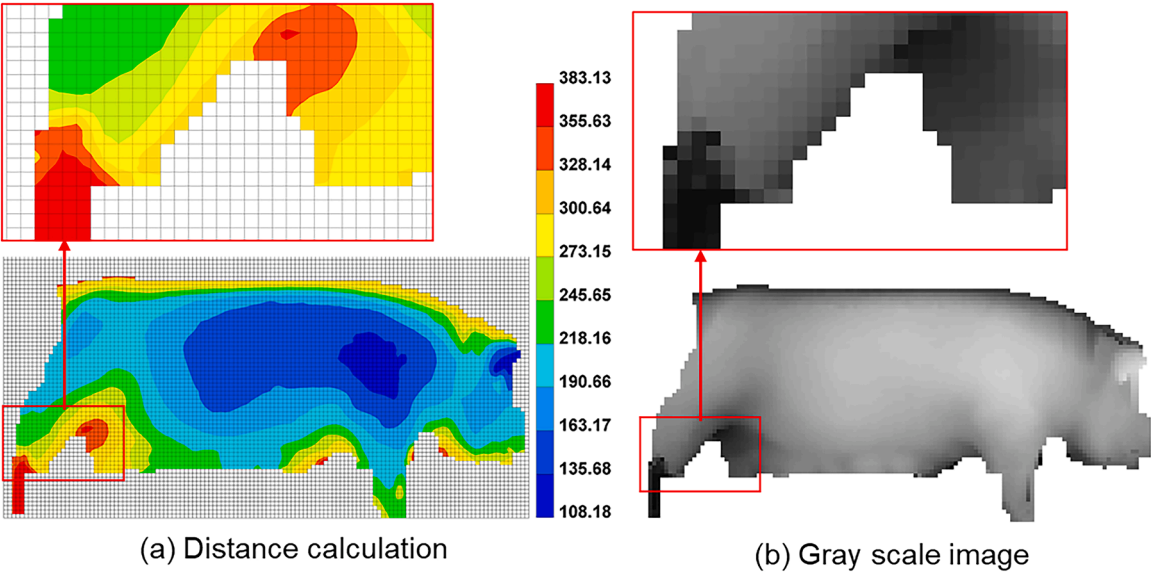


Fig. 11. Creating a grayscale image.

most convexly protruding sections are brightly represented. The brightest section has a value of 0.27, and darkest sections had value of 0.96. The boundary of the pig is represented by darker tones, which effectively distinguish it from the margin.

Fig. 12 shows the image transformation applied to the pig points of different weights. Fig. 12(a) shows a case in which the actual weight was 195 kg and relatively complete data were acquired for the entire side. Fig. 12(b) shows a weight of 249 kg, where the lack of data in the head region is reflected as missing areas in the transformed image. Fig. 12(c) shows a pig weighing 303 kg, which was among the heaviest specimens measured in this study. In all three cases, the body portion between the hind and front legs was fully transformed, and the larger pigs occupied a wider area in the image. Thus, the transformed image can be used to estimate the size of a pig. Fig. 13 shows the transformed images of various pigs measured at different positions. There were images transformed from fully measured areas and images transformed from measurements taken almost exclusively from the torso area. The center of the image changed depending on the composition of the points.

3.2.3. Generation of an additional image

The XY plane transformation was performed using the normal plane and axis of the cylinder. Moreover, the center value of all points was at the center of the image. Depending on the differences in these values, the position of the reference plane may vary, resulting in different distances for the points in each element. In other words, different point densities and distributions can generate different images. The points from the hind legs of the pig to its front legs were considered important data, the peripheral points were incrementally removed, and the internal point density was slightly modified.

Additional images were generated by removing some points at the boundary and randomly excluding the internal points. Specifically, the points of the three elements at the boundary area were removed, and within the internal element, 2%–3% of the points were randomly eliminated. The order in which the points in the elements were deleted

was implemented using three methods, each generating 30 images. In the first case, the lower and upper parts of the boundary were sequentially removed. In the second case, the left and right parts of the boundary were sequentially removed. In the final case, the two cases were applied in parallel. In the final image for each case, points corresponding to 90 elements were removed from the boundary area compared with the initial points, and in each image, 2–3% of the internal points were randomly removed compared with the initial points. Thus, 90 images were created for each set of point-cloud data.

Fig. 14(a) shows the image generated from the original points, whereas Fig. 14(b) and (c) show the final images generated from the points for each removal case. Fig. 14(b) shows a case in which points corresponding to 30 elements were removed in the vertical direction. There were 94,052 points, and the range of grayscale values is 0.26 to 0.97. Fig. 14(c) shows the case of removal in the horizontal direction, with grayscale values from 0.35 to 0.94. Fig. 14(d) shows the case where points were simultaneously removed in both the vertical and horizontal directions, with grayscale values from 0.33 to 0.96. Additional images generated through this process exhibited slight differences in the range of grayscale values and the pixels containing these values.

3.3. CNN-based weight estimation

A CNN comprises a convolution layer that extracts features from the input data, a pooling layer that reduces the size of the data while preserving key features, and a fully connected layer. This was used for the classification and regression. In this study, a regression model was implemented to estimate the weight of pigs using the generated images.

The CNN algorithm was implemented using TensorFlow and Keras libraries. The model was trained on a system with an Intel Core i7 3.4 GHz CPU, 32GB of memory, and an NVIDIA V100 GPU server with 16GB of memory. 132,050 images were used for training out of the total 146,750 data points, and 14,700 images were used for testing (Fig. 15). A total of 26,410 images, accounting for 20% of the training data, were

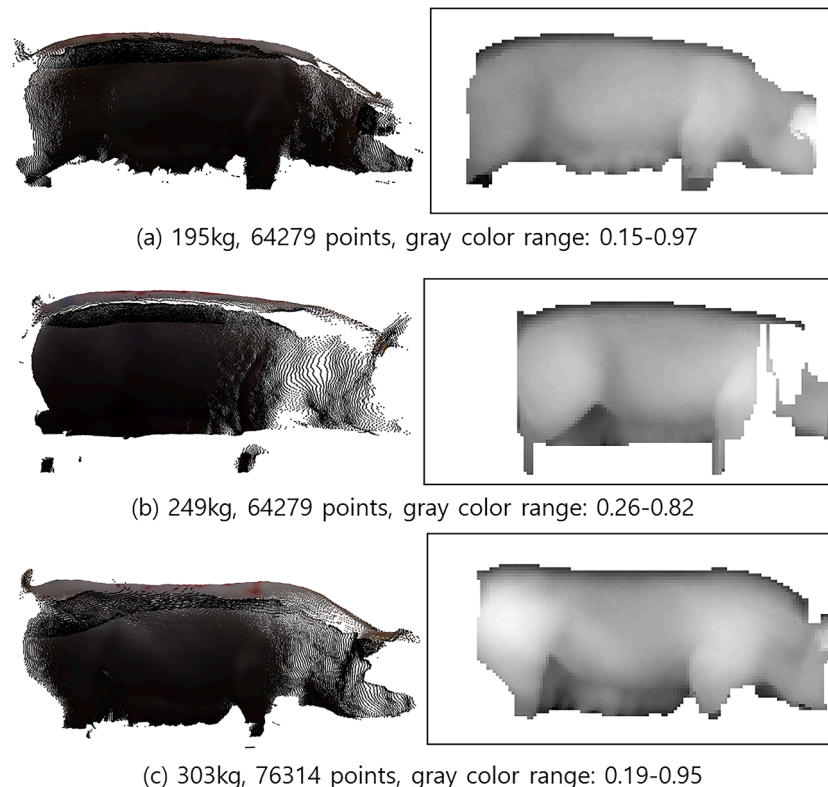


Fig. 12. Greyscale images generated for pigs of various weights.

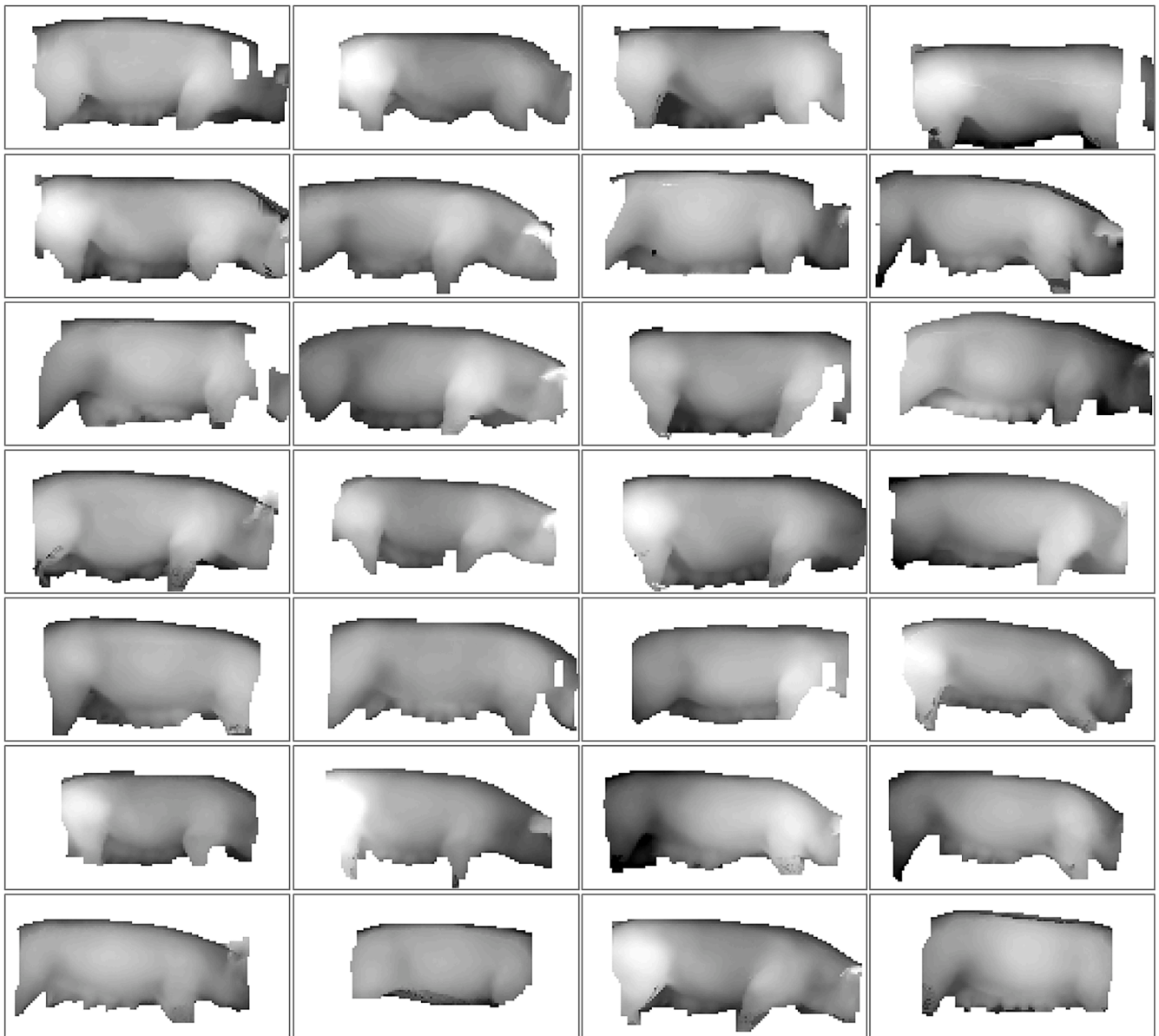


Fig. 13. Greyscale images generated for pigs in different positions.

used for validation. The additional images generated from the single point cloud were similar. Therefore, for testing, only images generated from the point-cloud data that were not used for training were used.

In this study, three well-known CNN models such as VGG16, DenseNet121, and EfficientNet B0 were applied using the Keras API to implement the neural network model (Fig. 16). CNN models have been widely used in livestock weight estimation studies, showing high performance by extracting meaningful characteristics from images obtained with RGB or RGB-D cameras [18,24,30–32]. Additionally, the three selected CNN models have unique advantages that make them suitable for addressing the pig weight estimation problem.

The first model, VGG16, is a simple deep structure model consisting of 16 layers and is one of the representative models in early computer vision research. It consists of 13 convolutional layers and 3 fully connected layers, but in this experiment, only 13 convolutional layers and one fully connected layer were used. VGG16 has a basic CNN structure, making it suitable for use as a comparative model. It minimizes the number of parameters using small-sized kernels, making it easier to train, while its deep structure provides excellent performance [44].

The second model, DenseNet, is a convolutional neural network

architecture that implements efficient deep learning models through dense connectivity. In DenseNet, the input to each layer is connected to the feature maps of all previous layers, enabling efficient information sharing. This addresses the issue of vanishing gradients and prevents the loss of information from the feature maps of earlier layers. Since the feature maps from previous layers continue to be used as input, DenseNet can be trained with fewer parameters compared to traditional CNN models, and it is less prone to overfitting, even with small datasets [45]. In this study, DenseNet121, which consists of 121 layers and is suitable for small images, was used.

The final model, EfficientNet, is known for its high performance in terms of the number of parameters. EfficientNet was designed to achieve a balance between model size and efficiency. It optimizes performance by adjusting the depth, width, and resolution of the network. Consequently, it achieves a high classification accuracy with a smaller model size [46]. EfficientNet has different sizes and complexities ranging from B0 to B7. In this study, a smaller and lighter EfficientNet B0 model was applied.

Moreover, instead of using models like PointNet, the CNN-based approach was chosen to better align with the characteristics of the

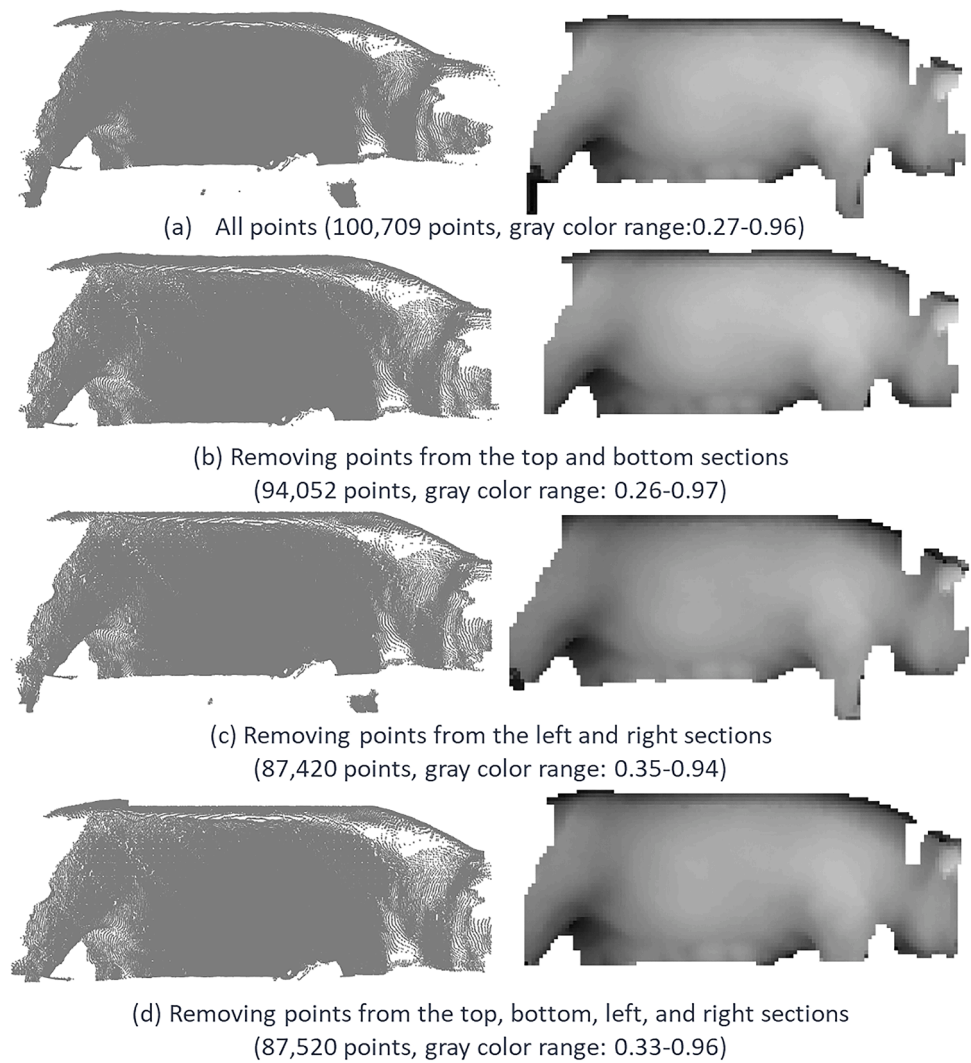


Fig. 14. Creating additional images.

Original set (146,750)		
Training set (132,050)		Test set (14,700)
Real training set (105,640)	Validation set (26,410)	Test set (14,700)

Fig. 15. Dataset partitioning.

data and the technical requirements of this study. When using point cloud data directly, accurate weight prediction becomes challenging due to the incompleteness and noise in the data collected while pigs are moving. Missing data frequently occurs in areas such as the head, tail, and legs, and point cloud data requires additional corrections during post-processing. Solving these issues would require more cameras and a separate scanning space, which is inefficient in a smart farming environment. Another reason for choosing the CNN-based approach is the advantage of using data augmentation techniques when converting point cloud data into 2D images for training. Data augmentation allows the creation of various training scenarios with a limited amount of data, improving the model's generalization performance. Thus, in this study,

the approach of converting point cloud data into 2D images and predicting pig weights using CNN models was selected. Model evaluation was performed using metrics such as the MAE (Eq. (4)), MAPE (Eq. (5)), root mean-squared error (RMSE) (Eq. (6)), and R2 (coefficient of determination) (Eq. (7)). Here, y_i is the actual value, $\hat{y}_i$ is the predicted value, $\bar{y}_i$ is the average of the actual values, and n is the total amount of data. All three models exhibited a MAPE within 2%, with DenseNet121 exhibiting the best performance at 1.35% (Table 1). Fig. 17 shows the graphs of the actual and predicted values for each model; large errors occurred for some of the test data. Fig. 18 shows the MAPE ratio for DenseNet121, which demonstrates the best performance. In total, 66.3% of the test data had an error of less than 1%, whereas

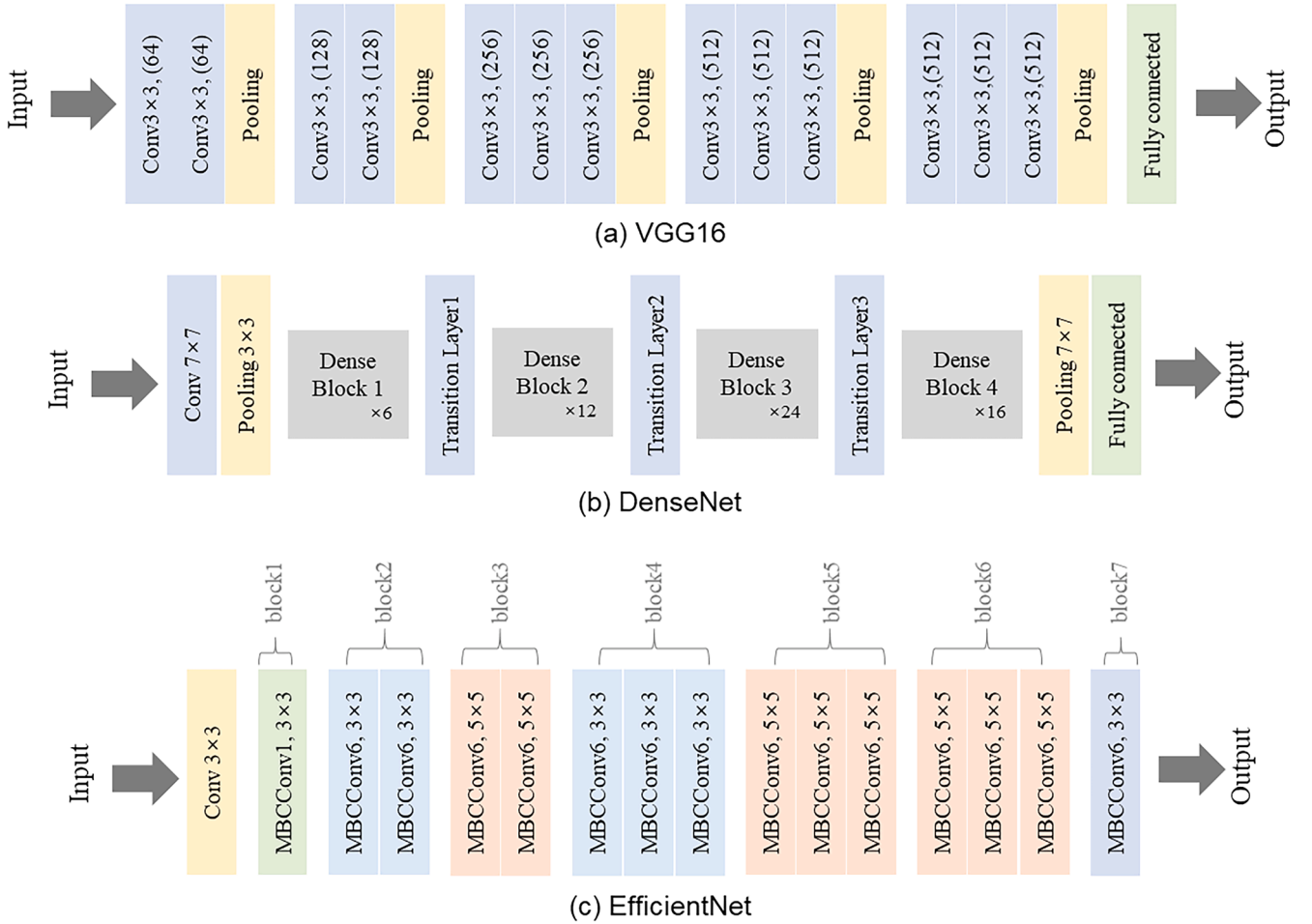


Fig. 16. CNN models used in training.

Table 1

Wt prediction performance in terms of MAE, MAPE, RMSE, and R^2 .

Model	Training dataset				Test dataset			
	MAE	MAPE	RMSE	R^2	MAE	MAPE	RMSE	R^2
VGGNet16	0.7144	0.3178	0.9508	0.9994	3.4219	1.5473	7.3342	0.9620
DenseNet121	0.9749	0.4309	1.5845	0.9985	2.9714	1.3546	5.3192	0.9800
EfficientNet B0	0.9712	0.4144	1.2219	0.9991	3.1253	1.4184	6.0230	0.9743

1.1% of the data had an error exceeding 10%.

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (4)$$

$$MAPE = \frac{100\%}{n} \sum_{i=1}^n \frac{|y_i - \hat{y}_i|}{y_i} \quad (5)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n}} \quad (6)$$

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (7)$$

Figs. 17 and 18 present the prediction results for 14,700 test images graphically. A total of 90 images were generated for each measurement. Fig. 19 compares the actual values with the averages of the predicted values from the 90 images. Overall, it was confirmed that the weight

prediction was performed accurately.

Fig. 20(a) shows the errors for the 60 images generated from the two sets of measurement data, where the error was the greatest. These images were generated while gradually removing small amounts from the bottom and top parts of the boundary area. Fig. 20(b) and (c) show images corresponding to the parts circled in Fig. 20(a). In Case 1, there was a 21% error when nothing was removed, and an 8% error when some points were removed. In Case 2, there was a 1% error when nothing was removed. However, the error started to increase from the 13th image and did not increase further from the 18th image. Errors occurred depending on the area in which the image existed; such errors did not occur for the other test data. It is suspected that this phenomenon occurred because of the lack of training data.

Fig. 21 shows three data points continuously measured as the pig walks, which were cases where many errors occurred in a previous study. Fig. 21(a) corresponds to Case 1 in Fig. 20. It was highly bent, resulting in a large difference in the data between the left and right sides. In this study, there was no effect on the left or right side as a result of our implementation.

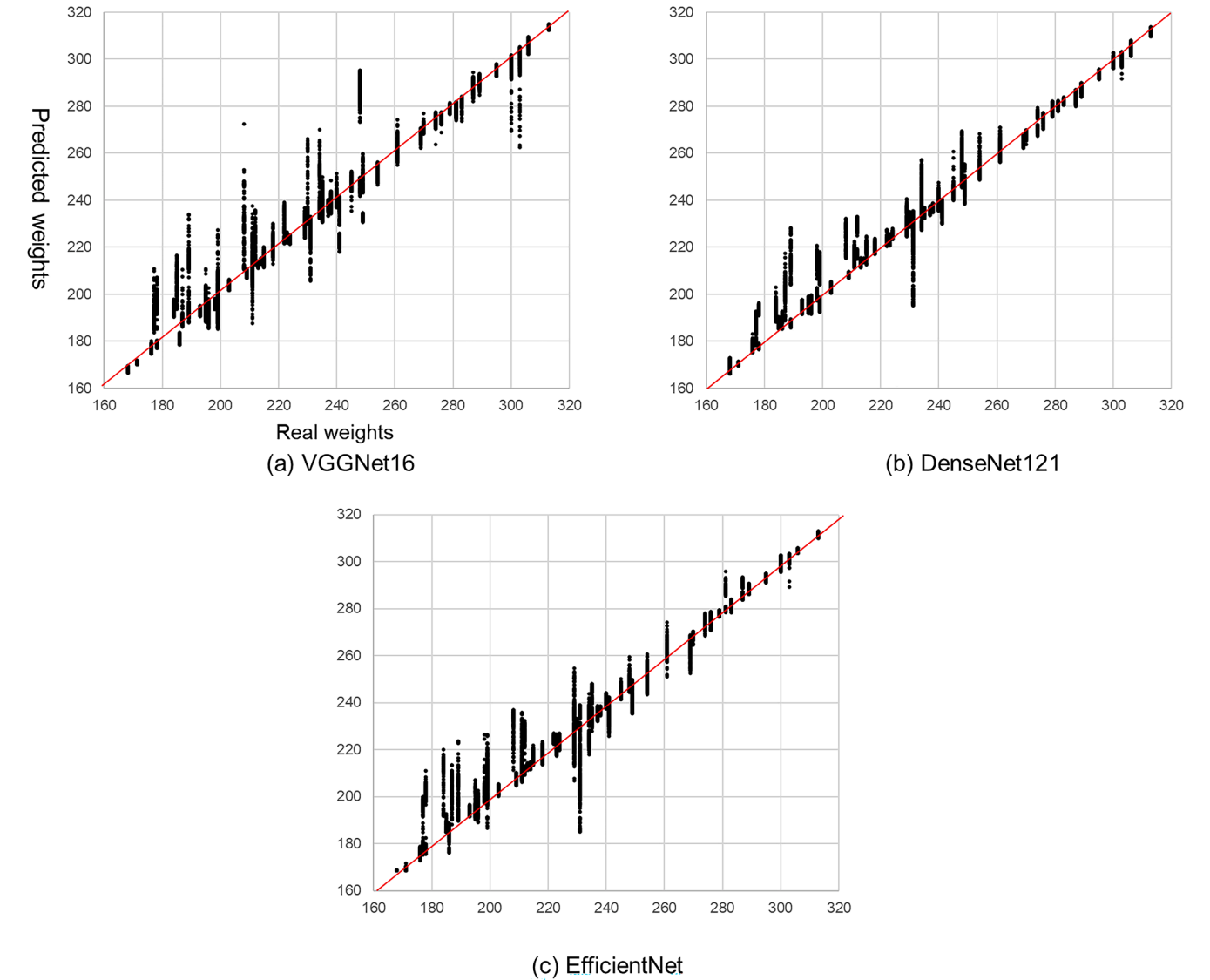


Fig. 17. Estimation results with test dataset.

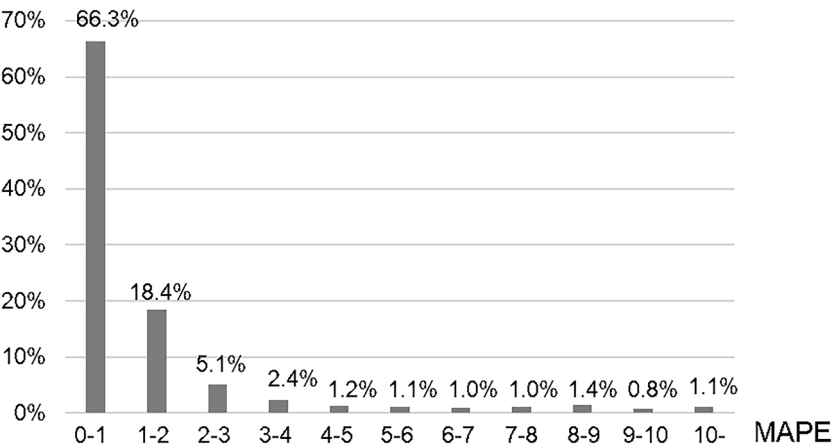


Fig. 18. Distribution of error rate (DenseNet121).

4. Discussion

In this study, a CNN-based method was proposed to estimate the

weight of pigs using point cloud data. The performance of three CNN models (VGG16, DenseNet121, and EfficientNet B0) was analyzed, and each model showed the performance described in Table 1. Upon

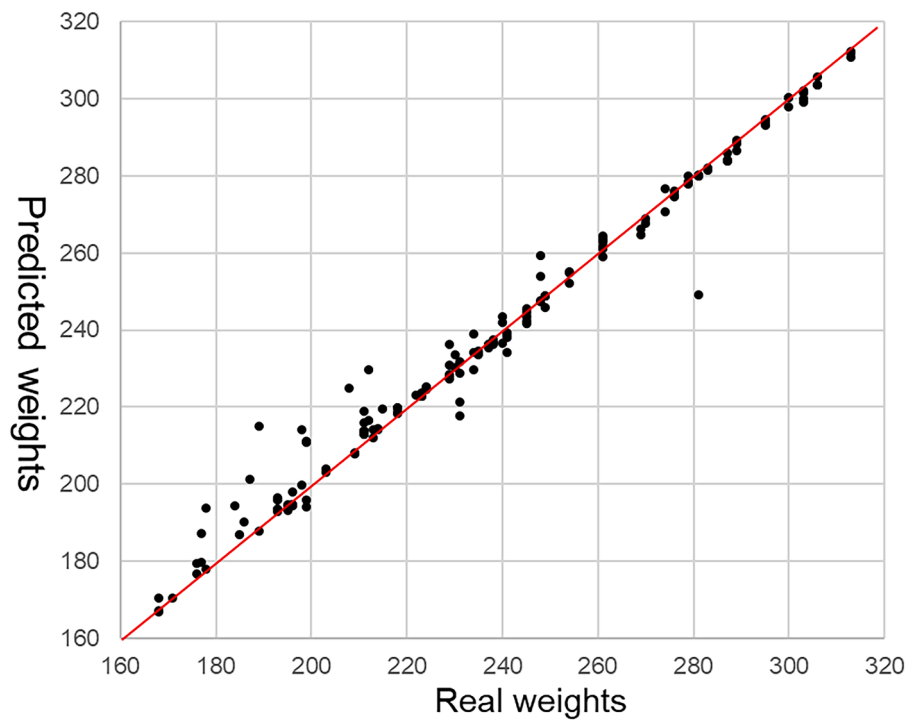
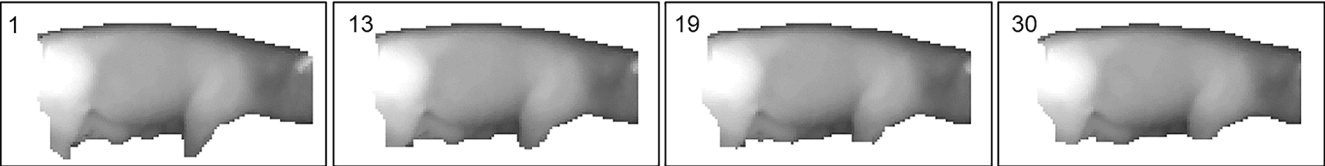
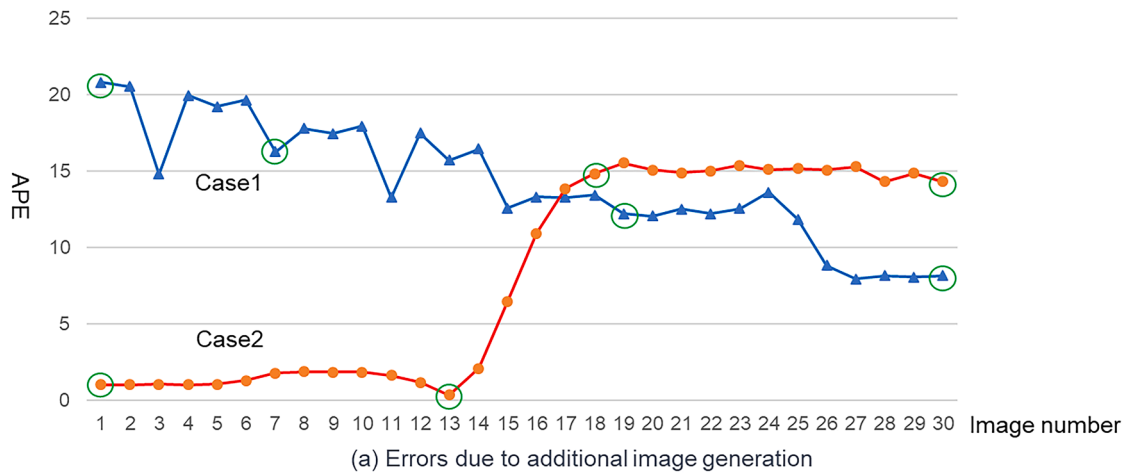
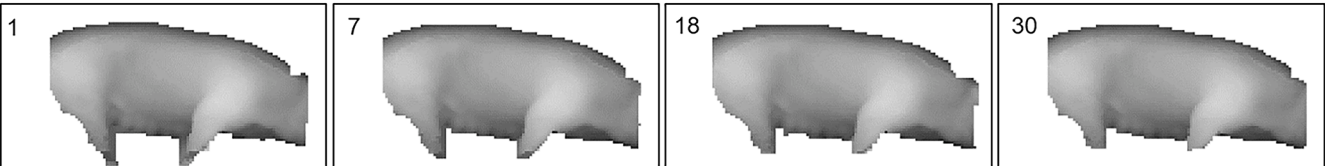


Fig. 19. Estimation results with test dataset (average value, DenseNet121).



(b) Images of Case1



(c) Images of Case2

Fig. 20. Cases where significant errors occurred.

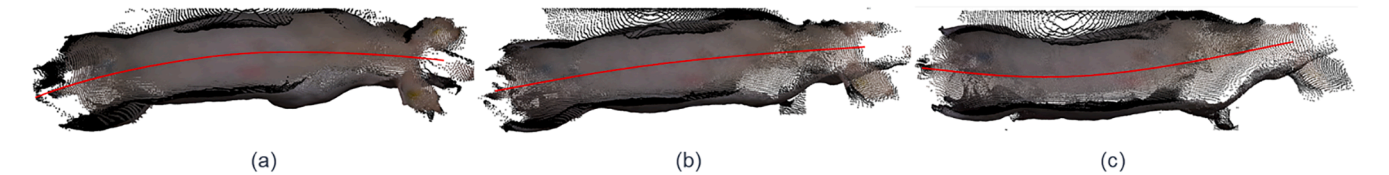


Fig. 21. Continuously measured data for the pig in Case 1.

analyzing the results, VGG16 recorded the lowest MAE, MAPE, and RMSE, as well as the highest R^2 value on the training dataset, demonstrating excellent training performance. However, its performance significantly declined on the test dataset, suggesting the possibility of overfitting. On the other hand, DenseNet121 showed the highest error on the training dataset but recorded the lowest error on the test dataset. This can be attributed to DenseNet's unique structural feature of connecting layers, which allows efficient reuse of features, resulting in high generalization performance. Additionally, since the dataset used in this study is relatively small, the inter-layer connectivity of DenseNet likely minimized information loss and helped address the vanishing gradient problem commonly encountered in deep learning models. These characteristics likely enabled DenseNet to secure better generalization performance than the other models, leading to stable prediction performance on the test data.

The data used in this study focused on obtaining complete data for half of the pig's body by dividing the entire body in half. This approach contributed to improving the accuracy of weight estimation, but obtaining complete side data for animals in clustered or moving conditions remains a challenge. To enhance the efficiency of weight estimation, it is necessary to apply point clouds obtained from a single camera, but this may lead to variations in the animal's center and orientation. Currently, the center of the image is set based on the center point of the entire points used, but replacing this with the center point of the body could better reflect the characteristics. The center of the body can be determined based on the front and rear legs. Since the body can be judged by using the front and rear legs, this approach should be considered in future studies. Moreover, even when the maximum-sized pig is applied, there is still some margin in the converted image. To better capture the characteristics of the pig in smaller images, improvements in pixel size are necessary. In this study, an image size of 128×64 was used, but applying a higher resolution is expected to yield more accurate results, which would require the acquisition of more data. Although this study was limited to pigs, there are plans to apply the method to other livestock such as cattle. Since the size of the livestock differs, appropriate pixel settings, as well as the image's center and orientation, need to be determined. For cattle, only part of the side might be captured in a single image, and combining images of different side parts might be necessary to address this issue.

Additionally, in a Windows 11 environment using a PC equipped with an i9-13900 K CPU, it took less than 0.05 s to generate a single grayscale image from point cloud data composed of hundreds of

thousands of points. Testing 14,700 grayscale images took approximately 90 s, with an average of 0.0061 s per image. This fast data generation and testing capability offer great advantages in processing and analyzing large amounts of data. It also opens up the possibility of improving accuracy by removing outliers and using the average value from multiple images.

Finally, Table 2 compares the results with previous studies. Most studies applied to pigs have focused on growing-finishing pigs, estimating weight using side or top-view images, but despite the simple shapes of the livestock, accuracy has not been high. The method proposed by Dohmen et al. [41] shows relatively high accuracy, but it was applied to young cattle, and the 2D images lacked depth information, requiring correction. The results of this study are significant in that point cloud data were used, reflecting both size and depth information, which enables accurate weight estimation automatically.

5. Conclusion

Predicting the weight of moving livestock is challenging because of their changing postures. In the case of sows, a comprehensive acquisition of data encompassing the entire body is necessary for accurate weight prediction. This typically involves the installation of multiple cameras and the use of supporting structures. This study demonstrates the potential of a CNN-based method for weight prediction of walking breed sows using partial measurement data. The method converts point cloud data representing half of the pig's body into 2D grayscale images, which are then used as input for CNN models to predict weight. This approach simplifies the camera setup requirements and enables accurate predictions even in dynamic environments where the animals' postures are constantly changing, effectively capturing data in such conditions.

The results showed that all three CNN models achieved high accuracy, with error rates below 2%, and DenseNet121 performed the best, with a MAPE of 1.35%. This suggests that CNN models can effectively compensate for variations due to posture changes, and the methodology of using point cloud data and grayscale images offers a valid alternative to traditional weight prediction methods. However, some of the data exhibited significant errors, and subtle differences in the images influenced the predictions. A total of 1630 measurement data points from 71 pigs were used, and it was judged that the phenomenon occurred owing to the insufficient amount of training data.

In the future, it will be important to secure more training data to improve the robustness of the model. Although this study focused on

Table 2
Comparison of the present study with previous research.

	Present study	Kwon and Mun, 2023 [26]	Liu et al., 2023 [32]	Ruchay el al, 2022 [18]	Dohmen et al., 2021 [41]	Meckbach et al., 2021 [30]
Animal	Breed sows	Breed sows	Growing-finishing pigs	Cows	Heifers	Growing-finishing pigs
Weight range	168 kg to 313kg	168 kg to 313kg	70 to 150kg	243 to 605kg	37 to 370kg	20 to 133kg
Data acquired	Point cloud	Point cloud	RGB images	Point cloud		RGB images
Measurement range	Side	Full	Side	Side	Top/Side	Top
Data conversion	Gray scale images	3D Meshes	RGB images	RGB images	RGB images	RGB images
Weight estimation	CNN	Deep neural network	CNN	CNN		CNN
Estimation error	MAPE: 1.35%	MAPE: 2.11%	MAPE: 2.82%	MAPE:8.4%	RSME: 20kg	MAPE:3.9%

breed sows, it also suggests the potential for expanding the methodology to other livestock species such as cattle, with the need for data collection and processing techniques tailored to the size and shape of each animal. The results of this study contribute to the advancement of automated weight prediction systems and are expected to play a significant role in improving the operational efficiency of the livestock industry.

Funding

This work was supported by the National Research Foundation of Korea (NRF) through the Basic Science Research Program, funded by the Korean government (Project No. 2021R1F1A1048172); the Industrial Technology Alchemist Project Program (No. RS-2024-00,419,010) funded by the Korean Ministry of Trade, Industry & Energy(MOTIE, Korea).

Ethics statement

Not applicable: This manuscript does not include human or animal research.

If this manuscript involves research on animals or humans, it is imperative to disclose all approval details.

CRediT authorship contribution statement

Kiyoun Kwon: Writing – original draft, Supervision, Software, Methodology, Funding acquisition, Conceptualization. **Jun Hwan Park:** Software, Validation, Writing – review & editing. **Ahram Park:** Software, Data curation. **Sangwoo Kim:** Software, Data curation. **Nojun Lee:** Software, Data curation. **Duhwan Mun:** Writing – review & editing, Writing – original draft, Validation, Funding acquisition.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Kiyoun Kwon reports financial support was provided by National Research Foundation of Korea.

Duhwan Mun reports financial support was provided by Korean Ministry of Trade, Industry and Energy.

Data availability

The authors do not have permission to share data.

References

- Brandl, E., Jørgensen, Determination of live weight of pigs from dimensions measured using image analysis, *Comput. Electron. Agric.* 15 (1) (1996) 57–72, [https://doi.org/10.1016/0168-1699\(96\)00003-8](https://doi.org/10.1016/0168-1699(96)00003-8).
- E.L. Wagner, P.J. Tyler, A comparison of weight estimation methods in adult horses, *J. Equine Vet. Sci.* 31 (12) (2011) 706–710, <https://doi.org/10.1016/j.jevs.2011.05.002>.
- M. Tschärke, T. Banhazi, Review of methods to determine weight and size of livestock from images, *Aust. J. Multi-Discip. Eng.* 10 (1) (2013) 1–17, <https://doi.org/10.1158/14488388.2013.11464860>.
- C.P. Schofield, J.A. Marchant, R.P. White, N. Brandl, M. Wilson, Monitoring pig growth using a prototype imaging system, *J. Agric. Eng. Res.* 72 (3) (1999) 205–210, <https://doi.org/10.1006/jaer.1998.0365>.
- N.R. Augspurger, M. Ellis, Weighing affects short-term feeding patterns of growing-finishing pigs, *Can. J. Anim. Sci.* 82 (3) (2002) 445–448, <https://doi.org/10.4141/A01-046>.
- R. Dingwell, M. Wallace, C. McLaren, C. Leslie, K. Leslie, An evaluation of two indirect methods of estimating body weight in holstein calves and heifers, *J. Dairy Sci.* 89 (10) (2006) 3992–3998, [https://doi.org/10.3168/jds.S0022-0302\(06\)72442-0](https://doi.org/10.3168/jds.S0022-0302(06)72442-0).
- T. Grandin, C. Shivley, How farm animals react and perceive stressful situations such as handling, restraint, and transport, *Animals* 5 (4) (2015) 1233–1251, <https://doi.org/10.3390/ani5040409>.
- L. Fautitano, S. Goumon, Transport of pigs to slaughter and associated handling, advances in pig welfare, woodhead publishing series in food science, *Technol. Nutr.* (2018) 261–293, <https://doi.org/10.1016/B978-0-08-101012-9.00009-5>.
- S. Bhoj, A. Tarafdar, A. Chauhan, M. Singh, G.K. Gaur, Image processing strategies for pig liveweight measurement: updates and challenges, *Comput. Electron. Agric.* 193 (2022) 106693, <https://doi.org/10.1016/j.compag.2022.106693>.
- R. Dohmen, C. Catal, Q. Liu, Computer vision-based weight estimation of livestock: a systematic literature review, *New Zealand J. Agric. Res.* 65 (2–3) (2022) 227–247, <https://doi.org/10.1080/00288233.2021.1876107>.
- Y. Zhao, Q. Xiao, J. Li, K. Tian, L. Yang, P. Shan, X. Lv, L. Li, Z. Zhan, Review on image-based animals weight weighing, *Comput. Electron. Agric.* 215 (2023) 108456, <https://doi.org/10.1016/j.compag.2023.108456>.
- M. Kashiha, C. Bahr, S. Ott, C.P. Moons, T.A. Niewold, F.O. Ödberg, D. Berckmans, Automatic weight estimation of individual pigs using image analysis, *Comput. Electron. Agric.* 107 (2014) 38–44, <https://doi.org/10.1016/j.compag.2014.06.003>.
- C. Shi, J. Zhang, G. Teng, Mobile measuring system based on LabVIEW for pig body components estimation in a large-scale farm, *Comput. Electron. Agric.* 156 (2019) 399–405, <https://doi.org/10.1016/j.compag.2018.11.042>.
- K. Jun, S.J. Kim, H.W. Ji, Estimating pig weights from images without constraint on posture and illumination, *Comput. Electron. Agric.* 153 (2018) 169–176, <https://doi.org/10.1016/j.compag.2018.08.006>.
- A. Pezzuolo, M. Guarino, L. Sartori, L.A. González, F. Marinello, On-barn pig weight estimation based on body measurements by a Kinect v1 depth camera, *Comput. Electron. Agric.* 148 (2018) 29–36, <https://doi.org/10.1016/j.compag.2018.03.003>.
- A. Pezzuolo, M. Guarino, L. Sartori, F. Marinello, A feasibility study on the use of a structured light depth-camera for three-dimensional body measurements of dairy cows in free-stall barns, *Sensors* 18 (2) (2018) 673, <https://doi.org/10.3390/s18020673>.
- I.C. Condotta, T.M. Brown-Brandl, K.O. Silva-Miranda, J.P. Stinn, Evaluation of a depth sensor for mass estimation of growing and finishing pigs, *Biosyst. Eng.* 173 (2018) 11–18, <https://doi.org/10.1016/j.biosystemseng.2018.03.002>.
- A. Ruchay, V. Kober, K. Dorofeev, V. Kolpakov, A. Gladkov, H. Guo, Live weight prediction of cattle based on deep regression of RGB-D images, *Agriculture* 12 (11) (2022) 1794, <https://doi.org/10.3390/agriculture12111794>.
- S. Shuai, Y. Ling, L. Shihao, Z. Haojie, T. Xuhong, L. Caixing, S. Aidong, L. Hanxing, Research on 3D surface reconstruction and body size measurement of pigs based on multi-view RGB-D cameras, *Comput. Electron. Agric.* 175 (2020) 105543, <https://doi.org/10.1016/j.compag.2020.105543>.
- I.C. Condotta, T.M. Brown-Brandl, R.V. Sousa, K.O. Silva-Miranda, Using an artificial neural network to predict pig mass from depth images, in: *Proceedings of the 10th International Livestock Environment Symposium (ILES X)*, American Society of Agricultural and Biological Engineers, 2018, <https://doi.org/10.13031/iles.18-043>.
- A.H. Nguyen, J.P. Holt, M.T. Knauer, V.A. Abner, E.J. Lobaton, S.N. Young, Towards rapid weight assessment of finishing pigs using a handheld, mobile RGB-D camera, *Biosyst. Eng.* 226 (2023) 155–168, <https://doi.org/10.1016/j.biosystemseng.2023.01.005>.
- K. Wang, H. Guo, Q. Ma, W. Su, L. Chen, D. Zhu, A portable and automatic Xtion-based measurement system for pig body size, *Comput. Electron. Agric.* 148 (2018) 291–298, <https://doi.org/10.1016/j.compag.2018.03.018>.
- Y. Le Cozler, C. Allain, A. Caillot, J.M. Delouard, L. Delattre, T. Luginbuhl, P. Faverdin, High-precision scanning system for complete 3D cow body shape imaging and analysis of morphological traits, *Comput. Electron. Agric.* 157 (2019) 447–453, <https://doi.org/10.1016/j.compag.2019.01.019>.
- L. Huang, H. Guo, Q. Rao, Z. Hou, S. Li, S. Qiu, H. Wang, Body dimension measurements of qinchuan cattle with transfer learning from LiDAR sensing, *Sensors* 19 (22) (2019) 5046, <https://doi.org/10.3390/s19225046>.
- K.Y. Kwon, D. Mun, Iterative offset-based method for reconstructing a mesh model from the point cloud of a pig, *Comput. Electron. Agric.* 198 (2022) 106996, <https://doi.org/10.1016/j.compag.2022.106996>.
- K.Y. Kwon, A. Park, H. Lee, D. Mun, Deep learning-based weight estimation using a fast-reconstructed mesh model from the point cloud of a pig, *Comput. Electron. Agric.* 210 (2023) 107903, <https://doi.org/10.1016/j.compag.2023.107903>.
- A. Rahagiyanto, M. Adhyatma, A review of morphometric measurements techniques on animals using digital image processing, *Food Agric. Sci. Polije Proc. Ser.* 3 (1) (2021) 67–72.
- Q. Yang, D. Xiao, A review of video-based pig behavior recognition, *Appl. Anim. Behav. Sci.* 233 (2020) 105146, <https://doi.org/10.1016/j.applanim.2020.105146>.
- K. Jun, S.J. Kim, H.W. Ji, Estimating pig weights from images without constraint on posture and illumination, *Comput. Electron. Agric.* 153 (2018) 169–176, <https://doi.org/10.1016/j.compag.2018.08.006>.
- C. Meckbach, V. Tiesmeyer, I. Traulsen, A promising approach towards precise animal weight monitoring using convolutional neural networks, *Comput. Electron. Agric.* 183 (2021) 106056, <https://doi.org/10.1016/j.compag.2021.106056>.
- H. He, Y. Qiao, X. Li, C. Chen, X. Zhang, Automatic weight measurement of pigs based on 3D images and regression network, *Comput. Electron. Agric.* 187 (2021) 106299, <https://doi.org/10.1016/j.compag.2021.106299>.
- J. Liu, D. Xiao, Y. Liu, Y. Huang, A. Pig, Mass estimation model based on deep learning without constraint, *Animals* 13 (8) (2023) 1376, <https://doi.org/10.3390/ani13081376>.
- A.J. Heinrichs, W.C. Losinger, Growth of Holstein dairy heifers in the United States, *J. Anim. Sci.* 76 (5) (1998) 1254–1260, <https://doi.org/10.2527/1998.7651254x>.
- A.J. Heinrichs, B. Lammers, Monitoring Dairy Heifer Growth, College of Agricultural Sciences, PennState, 1998.

- [35] T. Grandin, C. Shivley, How farm animals react and perceive stressful situations such as handling, restraint, and transport, *Animals* 5 (4) (2015) 1233–1251, <https://doi.org/10.3390/ani5040409>.
- [36] J.E. Morris, G.M. Cronin, R.D. Bush, Improving sheep production and welfare in extensive systems through precision sheep management, *Anim. Prod. Sci.* 52 (7) (2012) 665–670, <https://doi.org/10.1071/AN11097>.
- [37] L. Huang, S. Li, A. Zhu, X. Fan, C. Zhang, H. Wang, Non-contact body measurement for qinchuan cattle with LiDAR sensor, *Sensors* 18 (9) (2018) 3014, <https://doi.org/10.3390/s18093014>.
- [38] M.S. Al Ard Khanji, C. Llorente, M.V. Falceto, C. Bonastre, O. Mitjana, Tejedor, Using body measurements to estimate body weight in gilts, *Can. J. Anim. Sci.* 98 (2) (2018) 362–367, <https://doi.org/10.1139/cjas-2016-0232>.
- [39] C. Shi, G. Teng, Z. Li, An approach of pig weight estimation using binocular stereo system based on LabVIEW, *Comput. Electron. Agric.* 129 (2016) 37–43, <https://doi.org/10.1016/j.compag.2016.08.012>.
- [40] Z. Liu, J. Hua, H. Xue, H. Tian, Y. Chen, H. Liu, Body weight estimation for pigs based on 3d hybrid filter and convolutional neural network, *Sensors* 23 (18) (2023) 7730, <https://doi.org/10.3390/s23187730>.
- [41] R. Dohmen, C. Catal, Q. Liu, Image-based body mass prediction of heifers using deep neural networks, *Biosyst. Eng.* 204 (2021) 283–293, <https://doi.org/10.1016/j.biosystemseng.2021.02.001>.
- [42] Y. Bi, L.M. Campos, J. Wang, H. Yu, M.D. Hanigan, G. Morota, Depth video data-enabled predictions of longitudinal dairy cow body weight using thresholding and Mask R-CNN algorithms, *Smart Agric. Technol.* 6 (2023) 100352, <https://doi.org/10.1016/j.atech.2023.100352>.
- [43] M. Gjergji, V. de Moraes Weber, L.O.C. Silva, R. da Costa Gomes, T.L.A.C. De Araújo, H. Pistori, M. Alvarez, Deep learning techniques for beef cattle body weight prediction, in: *Proceedings of the International Joint Conference on Neural Networks (IJCNN)*, IEEE, 2020, pp. 1–8, <https://doi.org/10.1109/IJCNN48605.2020.9207624>. July.
- [44] K. Simonyan, A. Zisserman, Very deep convolutional networks for large-scale image recognition, in: *Proceedings of the 3rd International Conference on Learning Representations (ICLR)*, Computational and Biological Learning Society, 2015, pp. 1–14.
- [45] G. Huang, Z. Liu, L. Van Der Maaten, K.Q. Weinberger, Densely connected convolutional networks, in: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2017, pp. 4700–4708.
- [46] M. Tan, Q. Le, Efficientnet: rethinking model scaling for convolutional neural networks, in: *Proceedings of the 36th International Conference on Machine Learning*, 2019, pp. 6105–6114.