

Low-cost livestock sorting information management system based on deep learning



Yuanzhi Pan ^{a,b,*}, Yuzhen Zhang ^{a,c}, Xiaoping Wang ^d, Xiang Xiang Gao ^b, Zhongyu Hou ^{b,e,**}

^a Artificial Intelligence Lab, Zhenjiang Hongxiang Automation Technology Co. Ltd., Zhenjiang 212050, Jiangsu Province, China

^b School of Electronic Information and Electrical Engineering, Shanghai Jiao Tong University, 800 Dongchuan Road, Shanghai 200030, China

^c School of Innovation and Entrepreneurship, Jiangsu University, 301 Xuefu Road, Zhenjiang 212013, Jiangsu Province, China

^d Shanghai Picowave Technology Co. Ltd., Shanghai 201802, China

^e Guangzhou Institute of Advanced Technology, Chinese Academy of Science (GIAT), Guangzhou 511458, China

ARTICLE INFO

Article history:

Received 22 April 2023

Received in revised form 21 August 2023

Accepted 23 August 2023

Available online 24 August 2023

Keywords:

Pig monitoring

Image recognition

Sorting system

WGAN

ResNet

ABSTRACT

Modern pig farming leaves much to be desired in terms of efficiency, as these systems rely mainly on electromechanical controls and can only categorize pigs according to their weight. This method is not only inefficient but also escalates labor expenses and heightens the threat of zoonotic diseases. Furthermore, confining pigs in large groups can exacerbate the spread of infections and complicate the monitoring and care of ill pigs. This research executed an experiment to construct a deep-learning sorting mechanism, leveraging a dataset infused with pivotal metrics and breeding imagery gathered over 24 months. This research integrated a Kalman filter-based algorithm to augment the precision of the dynamic sorting operation. This research experiment unveiled a pioneering machine vision sorting system powered by deep learning, adept at handling live imagery for multi-faceted recognition objectives. The individual recognition model based on Residual Neural Network (ResNet) monitors livestock weight for sustained data forecasting, whereas the Wasserstein Generative Adversarial Nets (WGAN) image enhancement algorithm bolsters recognition in distinct settings, fortifying the model's resilience. Notably, system can pinpoint livestock exhibiting signs of potential illness via irregular body appearances and isolate them for safety. Experimental outcomes validate the superiority of this proposed system over traditional counterparts. It not only minimizes manual interventions and data upkeep expenses but also heightens the accuracy of livestock identification and optimizes data usage. These findings reflect an 89% success rate in livestock ID recognition, a 32% surge in obscured image recognition, a 95% leap in livestock categorization accuracy, and a remarkable 98% success rate in discerning images of unwell pigs. In essence, this research augments identification efficiency, curtails operational expenses, and provides enhanced tools for disease monitoring.

© 2023 The Authors. Publishing services by Elsevier B.V. on behalf of KeAi Communications Co., Ltd. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

1. Introduction

With large-scale production in the breeding industry, the use of artificial intelligence technology for automated breeding has become a trend. Livestock production in many parts of the world has become an industrialized breeding business (Moekti, 2020), commonly referred to as industrialized animal husbandry. With increasing cost of labor in China and rising instances of livestock disease caused by zoonotic diseases, larger-scale farms and enterprises have elevated their demand

for intelligent and automated technologies (Latino et al., 2020). Adopting new technologies, such as the Internet of Things and deep learning, provides more possibilities for the wider popularization of intelligent and automated breeding (Neethirajan, 2020). Farms have adopted various technologies in view of the scale and nature of their operation. Large pig farms have adopted chip implantation technology by implanting chips into pig ears to identify (Marsot et al., 2020) and track (Pandey et al., 2021) pigs. The development process of the sorting system is depicted in Fig. 1. Small pig farms generally employ statistical methods based on the use of traditional ear tag; however, this approach can result in pigs biting each other's ears (Calderón Diaz et al., 2018), and daily maintenance requires frequent manual intervention. Therefore, it is necessary to adopt a universally automated sorting system.

With the development of new technologies, including pattern recognition, machine learning, artificial intelligence, and image recognition, more solutions are now available for the efficient sorting of livestock.

* Corresponding author at: Zhenjiang Hongxiang Automation Technology Co. Ltd., Zhenjiang 212050, Jiangsu Province, China.

** Corresponding author at: School of Electronic Information and Electrical Engineering, Shanghai Jiao Tong University, 800 Dongchuan Road, Shanghai 200030, China.

E-mail addresses: yzpan@connect.hku.hk, yuanzhi_p@hotmail.com (Y. Pan), zhyhou@sjtu.edu.cn (Z. Hou).

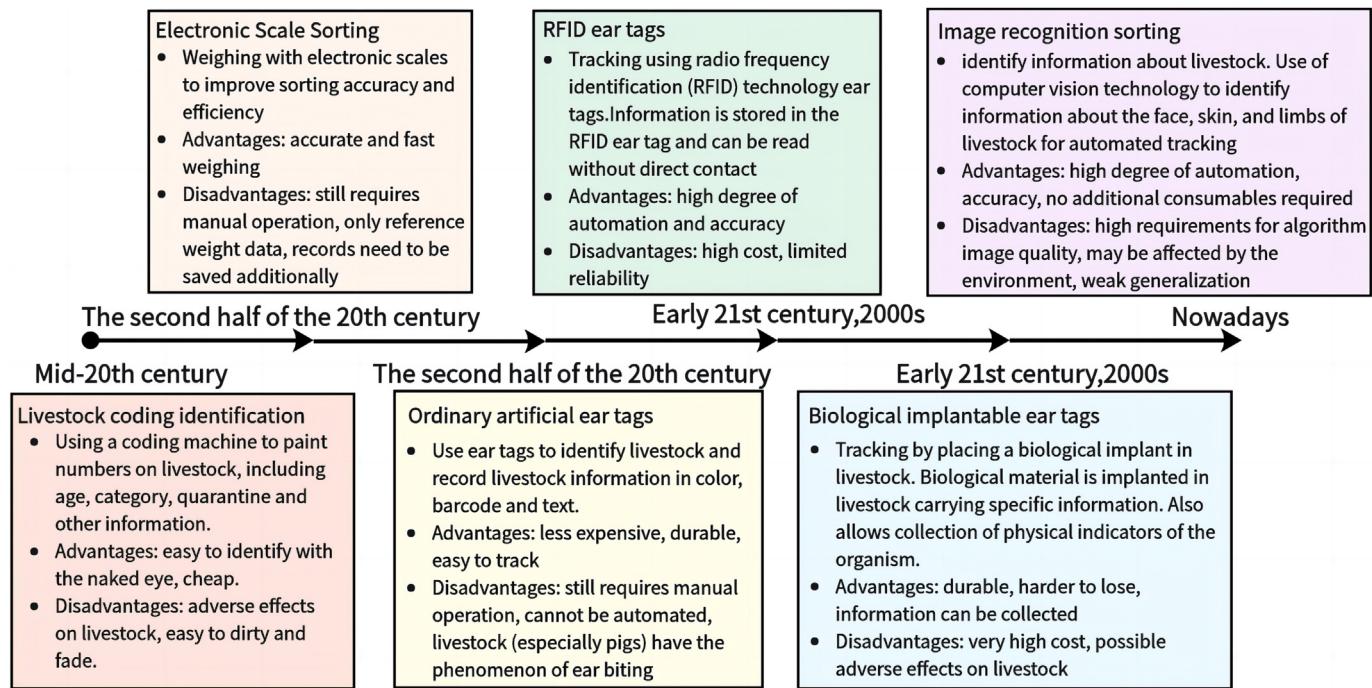


Fig. 1. Developmental history of livestock sorting.

By solving the many problems in previous sorting methods, labor costs can be effectively reduced, and production management efficiency can be improved. In our preliminary experimental study, it was found that the noise in the field environment, including the interference of visual acquisition and weighing equipment, affects the accuracy of the information acquisition, analysis, and judgment of the system. Ambient light and image blockage lead to insufficient or incorrect information contained in the image, thus affecting the detection accuracy. Fluctuations in the pressure of a moving hog as it passes through the weighing device can affect weight measurement, and thus the assessment of hog growth. Identification is highly dependent on the environment, and when an image is obscured, the identification accuracy is significantly reduced.

Image recognition, with its unique advantages of high automation, accuracy, and independence from additional consumables, has become a crucial focus of sorting systems. Complex image quality issues, particularly those involving glare spots or occlusions, necessitate auxiliary technologies associated with image completion to enhance image quality. WGAN-based image completion algorithms have been previously proposed to address these challenges. The original generative adversarial network (GAN) faced issues such as unstable training and uncontrollable generation; The algorithm was later improved upon by replacing the multilayer perceptron with a convolutional neural network (CNN), forming the deep convolutional GAN (DCGAN). While DCGAN offers increased stability in network training, it only mitigates some of the issues plaguing the GAN. The GAN is further improved by introducing the Wasserstein distribution distance GAN (WGAN), which provides a more stable training process (Radford et al., 2015).

Algorithms that enhance model transferability are required in response to the abundance of unlabeled images and the demand for robust generalization capabilities in real-world production environments. Transfer learning uses previously acquired knowledge (source domain with ample labeled data) to facilitate the solution of the current task. For instance, Miller et al. (2000) applied shared density to the digital transformation problem (Pan et al., 2011), and later designed a generative model based on the variational Bayesian framework (Miller et al., 2000). With the advent of CNNs in computer vision tasks,

deep-learning methods are being increasingly employed to solve few-shot learning problems.

Neural network algorithms have evolved to address technical challenges and improve the reliability and accuracy of models in complex and diverse application scenarios. Neural network fine-tuning, a type of deep transfer learning (Neethirajan, 2020), adapts pre-trained models with general feature extraction capabilities to target networks, adjusting the network structure based on specific training tasks. Fine-tuned model parameters are then trained on target datasets and applied to individual recognition tasks (Yosinski et al., 2014). Residual networks address the degradation problems in CNNs (Tzeng et al., 2015). In terms of application, Pan et al. (2022) proposed a computer vision-based recognition framework for distinguishing Neli-Ravi breeds from other buffalo breeds, achieving an accuracy of 93% using support vector machines and over 85% with recent variants. Kashiha et al. (2013) retained the pattern features of 10 pigs using Fourier descriptors with rotational and translational invariance and achieved 88.7% accuracy in pig pattern recognition.

Generative adversarial networks play a crucial role in generative models. The ResNet network is an effective deep-learning algorithm. Research on pigs encompasses various aspects including the processing of biological information (Wang et al., 2022), artificial intelligence breeding, and behavioral recognition. Furthermore, livestock research includes studies on mobile systems (Fuentes et al., 2022), digitalization of animal husbandry (Neethirajan, 2022), and emotional state of livestock (Wanga et al., 2015). Previous studies observe that the main characteristics of pigs are distributed in the skin, eyes, ears, nose, mouth, tail, and limbs, and their study can offer insights into the differences in pig characteristics. The main focus of this study is the extraction of these features and their application in recognition sorting. The feature descriptions are listed in the Table 1.

Based on the study of pig characteristics, we constructed a dataset containing the disease, breeding, growth, and fattening status of pig farms collected over two years of study to develop a deep-learning sorting system. The designed system processes field images with multiple feature-recognition tasks, such as disease and illness, sex and breeding, growth, and fattening.

Table 1

Morphological features of ternary pig.

Sr.	Characters	Description	Image
#			
1	Marking	The skin of pigs usually exhibits a uniform pink or light pink color. There may be some black or dark brown spots on the skin, typically distributed on the ears, face, and legs.	
2	Eye	The eyes are typically black or brown in color, with a dark appearance. The eyes are typically located on the sides of the head and slightly inclined forward.	
3	Ear	The ears are usually medium-sized and have a slightly drooping shape. The color of the ears can vary, but they are usually pink or light pink in color.	
4	Nose	The noses are usually flat and snout-like in shape. The nostrils are typically large and oval-shaped, allowing for efficient air intake.	
5	Tail	The tail is cylindrical in shape and tapers to a point. The tail is used for communication and can be moved to express different emotions such as excitement, happiness, or aggression.	
6	Feet	They typically have four toes on their front feet and three toes on their hind feet. The feet are relatively small in proportion to their body size. The feet are compact and rounded in shape, with short, sturdy toes.	
7	Healthy	Skin is smooth, elastic, and has well-arranged hair with normal sebum. Color varies from pink to light brown. Eyes are clear, secretion-free. Oral mucosa has normal color and saliva. Nose is moist without discharge. Limbs are strong, and joints lack redness, swelling, or inflammation.	
8	Diseased	Skin exhibits itching, redness, rashes, hair loss, and dullness. Eyes show congestion, edema, and increased secretions. Oral mucosa may have redness, ulcers, and bleeding; nose is dry or has discharge. Diseased pigs' limbs may have joint inflammation and unsteady gait, or limping. Weight gain is unstable, appetite is poor, and weight loss is rapid or emaciation occurs.	

This study identifies the shortcomings of deep-learning-based approaches applied in the livestock sorting field and aims to design a pig sorting system that encompasses information collection, data processing, optimized algorithm training, and relevant applications. In this study, a residual network algorithm model based on image completion that provides a low-cost and tractable solution for classification and prediction by integrating weight and image data is proposed.

2. Materials and methods

In this study, a sorting system that uses image acquisition to sort pigs is designed. Briefly, pigs enter the sorter from one end, and the sorter obtains signals to fix the pigs and collect data such as images and weights. After processing, the different gates are opened according to the sorting requirements for partitioning.

The sorting process:

- (1) The system first enables the pigs to pass in sequence through the infrared identification device at the entrance by means of a flexible position-limiting device that detects the imminent entry of the pigs and opens the front door. After the infrared detection device detects that the pig has fully entered, the front door is closed.

- (2) The system weighs the weight of the pig through an electronic tray and feeds the information back to the system. The pig images are collected through an image acquisition module.
- (3) The system detects the health of the pigs through image acquisition using a pre-trained model. Common pig diseases can cause abnormalities in the eyes, ears, nose, and skin. The system isolates sick pigs using image recognition, and the sick pigs are subsequently treated manually.
- (4) The system detects the collected images and compares different types and stages of pigs for sorting using a deep-learning model. The corresponding sorting gates are opened to achieve an orderly sorting of pig groups.
- (5) The ID weight data are managed to obtain the expected evaluation results.

2.1. System model

The proposed system consists of two parts: hardware and software. The hardware components include weight and image detection devices and the necessary mechanical parts, such as flexible limiters and sorting doors. The software encompasses image acquisition,

recognition, and sorting, which are implemented using deep-learning algorithms. A structural diagram of the system is shown in Fig. 2. In the case of pigs, livestock enter the sorting system from the mixed pens and are finally identified in the corresponding pens through the entry, identification, and sorting modules of the sorter, as shown in Fig. 2(A). The processes of sorting, modular device, and hardware organization are described in this study, and the star network structure is shown in Fig. 2(B). Modular channels, as shown in Fig. 2(C), have also been designed to achieve innovation from a traditional sorter, several of which together with the sorter, form an enclosed sorting area to achieve fully automated sorting. The hardware structure of the system is shown in Fig. 2(D). A flowchart of the system is shown in Fig. 2(E). The system has three components: detection and entry, data acquisition, and sorting and exit. Each part is designed in a modular manner. During the detection and entry processes, the system detects the entry of pigs through an infrared transceiver in front of the entrance gate and opens the gate. After the weighing module detects that a pig has entered the system, it closes the gate. The data acquisition process collects information such as pig weight and images.

In the experiment, a modular assembly device was designed, and a short-distance star communication network was built. It is controlled by a central control module, and each submodule has an independent control unit responsible for circuit control within the module. Each submodule is physically combined to form a sorting module consisting of at least one inlet, one detection, and several outlet modules. Similarly, connectors made of soft materials can be used to build pig passages with acoustic and optical devices to assist in the repelling of pigs and reduce human intervention; the components are assembled into the pigsty in a combination of soft and hard materials. Pigs can automatically complete the sorting process when driven by sound and light devices through guided passages.

The sorting process is divided into two parts. First, health detection was performed using the collected pig photographs. Combining the weight and visual information with a pre-trained binary classification model and decision-level fusion determines whether the pig is abnormal or sick, and sick pigs are sorted into isolation pens for workers to check and treat.

A fusion of the image neural network and weight-based classification results is used to improve the accuracy of the experiment and obtain the final target classification and recognition results through decision fusion. The ID and category of the pig are identified using the image, and its weight is measured using a weighing device. Sorting is performed by fusing the historical data of the pig with the new data according to its ID and evaluating the growth status of the pig and its growth level in the group. Simultaneously, image recognition detects whether there is a possibility of disease in the phenotype of the pig and sorts the pig to the diseased pen when it is detected as having a possible disease or abnormal weight.

2.1.1. Mechanical device

When the infrared detection device detects an obstacle and confirms that pigs are waiting to enter, the system controls the motor to open the entrance gate upward to guide the pigs in. After the electronic tray inside the device confirms that the pig has entered and detects no further obstacles in the infrared detection device, it closes the entrance gate and collects the weight data of the pig. After the image acquisition device has collected pictures and completed the classification work, the mechanical control device opens the corresponding partition exit gates, letting the pigs out of the sorting device. To sequentially guide the pigs into the sorter, a flexible position-limiting device is designed, as shown in Fig. 3. When the pig passes the threshold of ground pressure sensor B, the flexible limit device approaches and restricts the individual passing

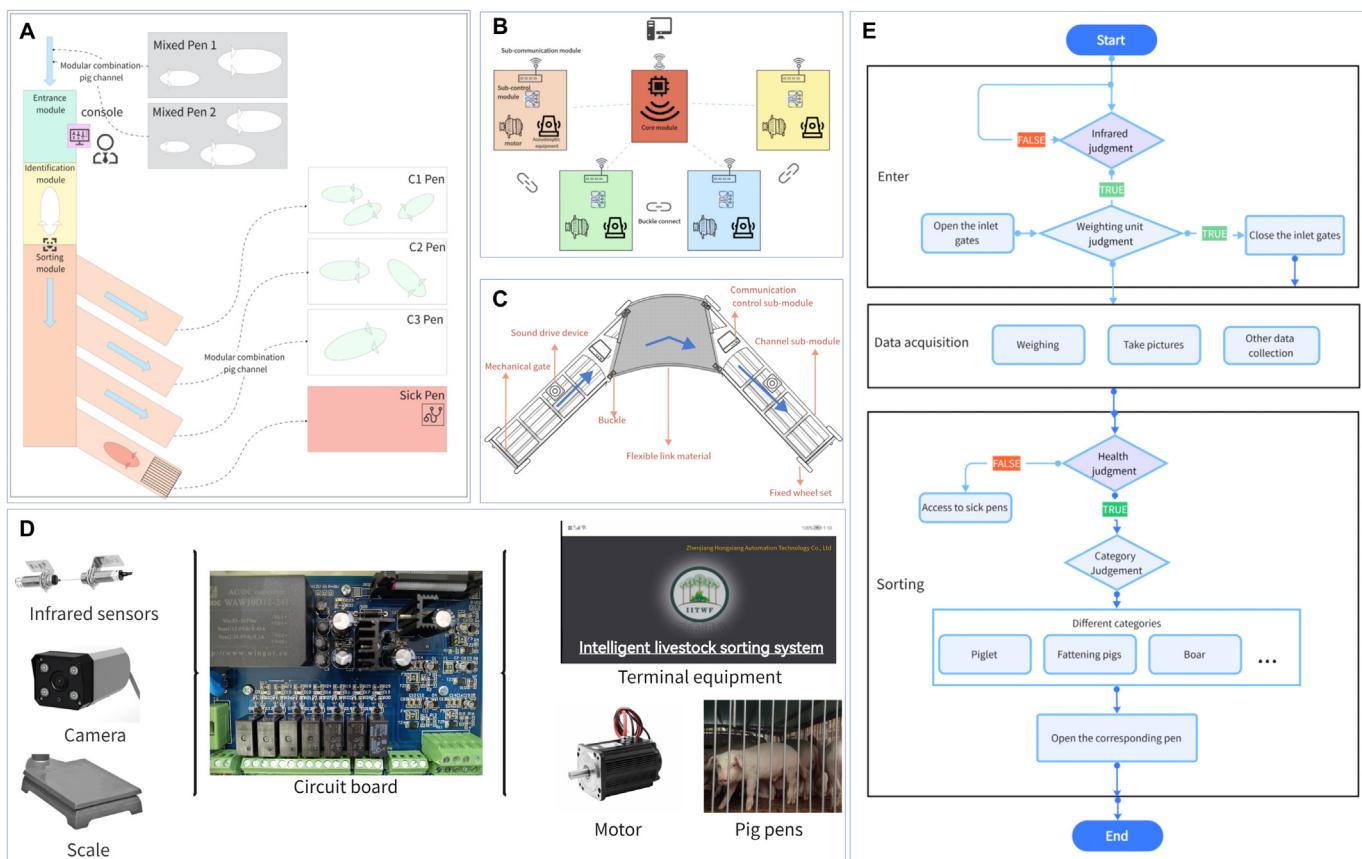


Fig. 2. (A) Automatic sorting systems in pig houses. (B) Star network structure. (C) Schematic of the module components. (D) Schematic of the system control. (E) System flow chart.

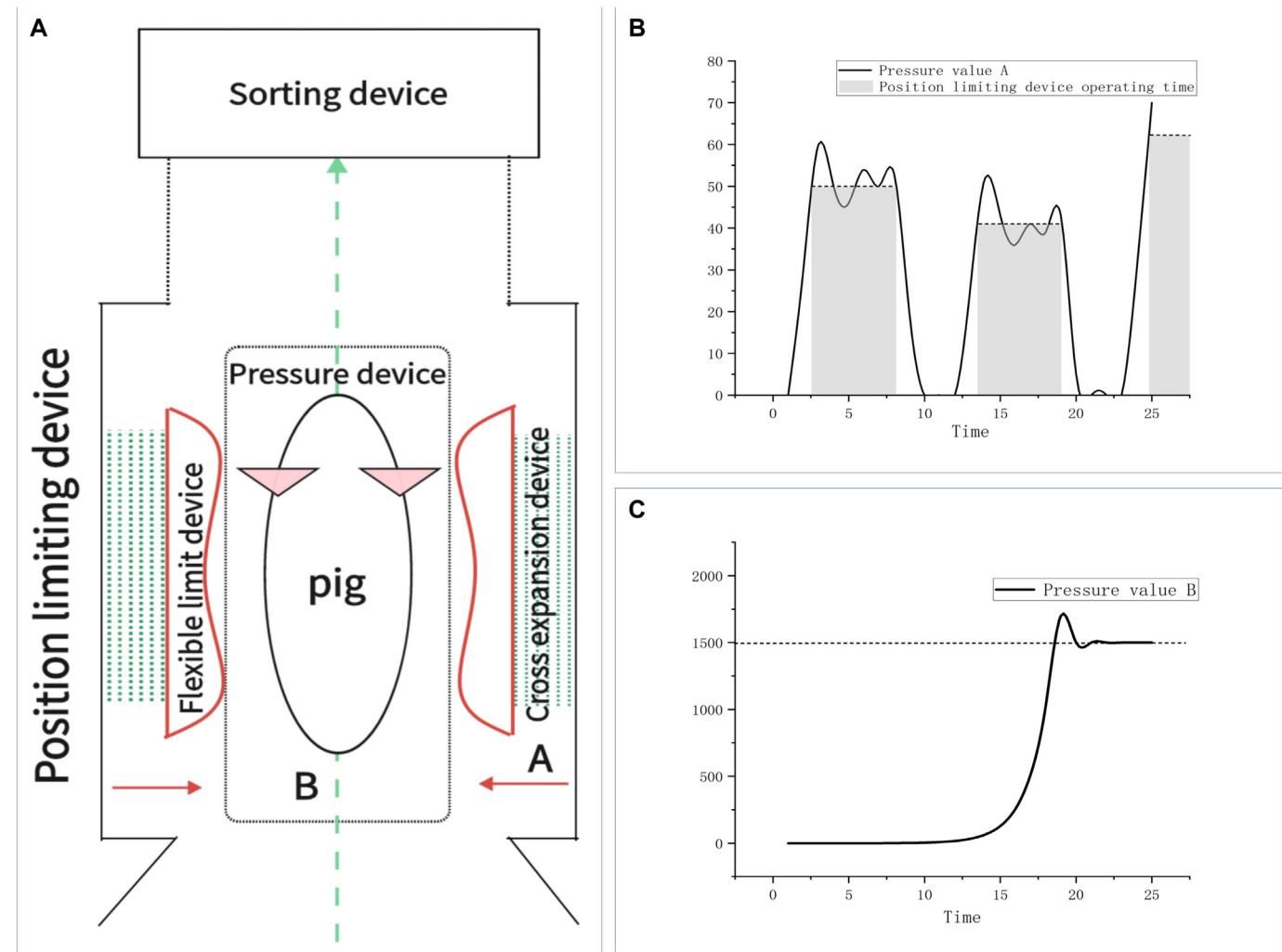


Fig. 3. (A) Position-limiting device. (B) Operation of pressure sensor B. (C) Operation of pressure sensor A.

of the pig. Side pressure sensor A detects contact and stops displacement; the pressure change curves of sensors A and B are shown in Figs. 3(C) and (B), respectively.

To achieve the sequential entry of pigs into the sorter, the position-limiting device can adjust the limiting space for different pigs. The entrance and detection modules are separated by a mechanically controlled gate that opens outward. The exit gate of the detection module opens into various branch channels. After the central processing module makes a sorting decision, the controller controls the opening of the gates on the corresponding sorting lanes via a star network communication to guide the pigs and complete the fast sorting.

2.1.2. Image acquisition device

This study considers the impact of the spectral characteristics of different light sources on the image acquisition devices in terms of factors such as cost and lifespan (Yin et al., 2012). Incandescent lamps are the most common light source, producing large amounts of infrared energy, and are inexpensive. They can also extend their usage time by operating at low voltages, but at the same time, they have the problem of low luminous efficiency (Dong et al., 2015); halogen lamps have a long service life and do not distort color but generate more heat; high-frequency fluorescent lamps generate less heat and have a long service life, but their color rendering is unsatisfactory (He et al., 2016a, 2016b); LED lamps produce less heat and have good monochromaticity. They achieve all colors in the visible light band with low power consumption.

When facing moving objects, they exhibit good shock and impact resistance (He et al., 2017). Therefore, LED lights are used as the visual light source (Kwon and Casebolt, 2006). Therefore, this experiment was designed using the image acquisition device, as shown in Figs. 4(A) and (B).

The LED fill light strip is located on both sides of the inner wall of the sorting device at a height of 1050 mm from the ground and acts as an effective supplementary light source. Two cameras are located 1300 mm above the ground on the central axis of the sorting device and 350 mm above the ground on the exit door. When entering, the camera is turned on after the gate is closed. Pictures of the front of the head of the pig, including the ears, nose, eyes, and mouth, and pictures of the neck and back from above are collected for sorting. The cameras are centered and located 1200 and 400 mm above the pigpen. To verify the robustness of the model, the light sources of the acquired images include natural, indoor, LED, and incandescent lights, as shown in Fig. 4(C). It was found that the age range of 50–150 days is a fast-growing stage for pigs with significant research value. In addition, the classification of pigs according to their development can better evaluate their growth.

Common pig diseases often present symptoms in the ears, nose, eyes, mouth, and skin (Ouyang and Ren, 2023), as shown in Fig. 5(A). Images acquired at important locations can effectively extract essential information (Robbins et al., 2014).

Images were collected over a period of three years from farms in Shandong Province, China. The image and weight data were tracked

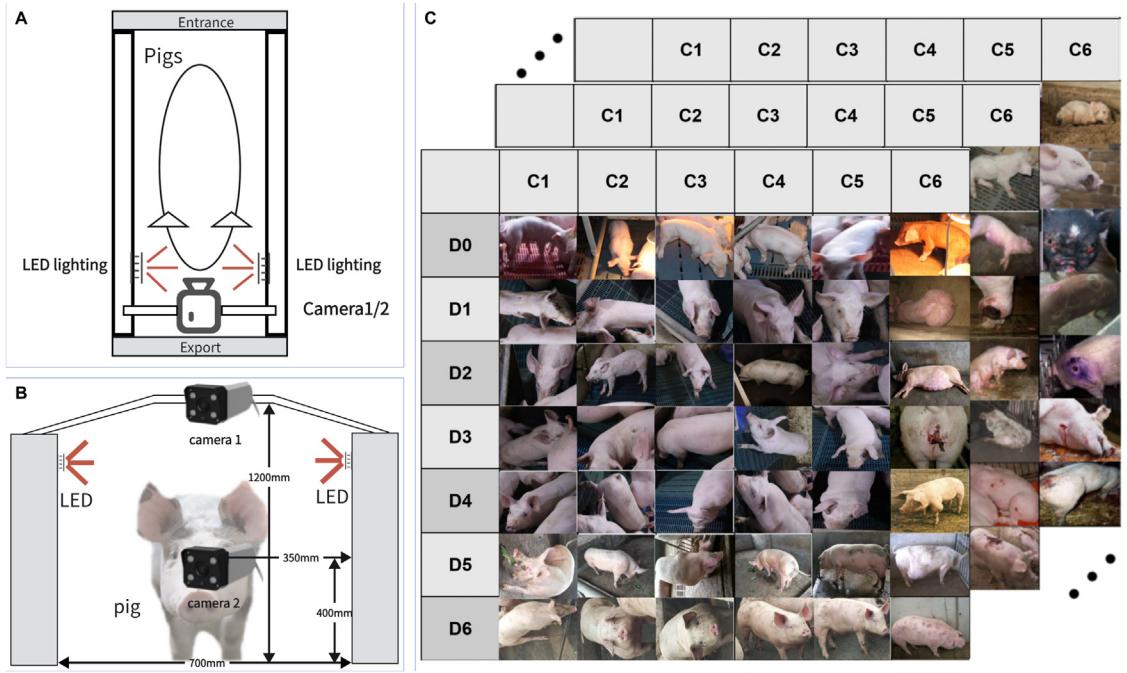


Fig. 4. (A) Top view of the image acquisition device. (B) Front view of the image acquisition device. (C) Acquisition of collated matrix images. C series from C1 to C6 are: stunted, poorly developed, normally developed, well developed, extremely well developed, and disease state. D series from D0 to D6 are: nursery pigs, 50 to 70 days old, 70 to 90 days old, 90 to 110 days old, 110 to 130 days old, 130 to 150 days old, and 150 days old to farrowing.

and summarized as statistics, as shown in Figs. 5(C). The age groups of the pigs used for data collection included nursery pigs and pigs aged 50 days to farrowing. The species were mainly Chinese ternary cross-bred pigs, including gestating sows and male breeding pigs. Image enhancement was also used to expand the dataset, including arbitrarily rotated images, images enhanced by Gaussian blur, and arbitrarily captured images, as shown in Fig. 5(B). Owing to the relative lack and limited variety of image samples collected from the sick pigs, sick pig images were expanded using an internet dataset.

2.1.3. Automatic weighing device

A voltage-type weighing sensor was used as the weighing device. Because the irregular movement of live pigs entering the device interferes with the measurement of body weight, Kalman filtering is used to predict the true weight of the live pigs. The voltage-type weighing sensor obtains real-time pressure and transfers it to the system. The system uses a Kalman filtering algorithm to achieve dynamic weighting. The process is illustrated in Fig. 6(A), and the operation of the Kalman filter is expressed as:

$$X_k = AX_{k-1} + BU_{k-1} \quad (1)$$

$$P_k = AP_{k-1}A^T + Q \quad (2)$$

Here, X_k is the state of the system at time k ; A is the state transition matrix related to the studied system; B is the input control matrix; U , an error matrix, is the external effect on the system; and Q is the predicted noise covariance matrix (Kristensen et al., 2012). Eq. (1) represents the prediction of the state, wherein X_{k-1} estimates the state at time k by recursively using the state at time $k - 1$. Eq. (2) represents the prediction of the error, wherein P_{k-1} uses the covariance at time $k-1$ to recursively obtain the covariance estimate at time k .

$$K_k = P_k H^T \left(H P_k H^T + R \right)^{-1} \quad (3)$$

$$X_k = X_k + K_k (Z_k - H X_k) \quad (4)$$

$$P_k = (I - K_k H) P_{k-1} \quad (5)$$

Here, K_k represents the Kalman gain, H is the observation matrix, R is the measurement noise covariance matrix, and Z_k is the observation value at time k . The Kalman gain can be calculated using Eq. (3). The corrected state and output of Kalman filtering are described by Eq. (4). The updated covariance matrix is expressed in Eq. (5). As shown in Fig. 6(B), the effect of the Kalman filtering algorithm is significantly better than that of the average value filtering algorithm. The Kalman filtering results significantly improve the sorting effect of the system.

The experiments were analyzed and evaluated using simulation tests, and the simulations were conducted in the range according to the actual weight. Let the state transfer matrix A be set as a unit array. The system does not influence the external input; U is zero, and H is a unit matrix. The gain matrix K does not need to be initialized, and the error matrix P is initialized as a zero matrix. Q and R are the predicted and observed state covariance matrices, respectively, and Q and R are set to 0.000015 and 0.008, respectively (Wang et al., 2017). A set of typical waveform graphs that can be considered as irregular motion were collected, leading to fluctuations in real-time weight information. Referring to the actual scenario using different data between 20 and 100 kg for the experiment, different sets of 20 data points were collected to calculate and compare the average and Kalman filtering errors.

The experimental results show that after the original weighing float, as shown in Fig. 6(C), is Kalman filtered, the weighing error is maintained within $\pm 0.7\%$, as shown in Fig. 6(D), which is clearly better than the average filter result and can meet the weighing requirements of the system.

2.2. Algorithm model

In this study, the backbone network of the algorithmic model is ResNet, as shown in Fig. 7(A). For occluded images, a WGAN-based image completion algorithm was employed to fill in the occlusions, thereby enhancing the overall recognition accuracy. Pre-trained models

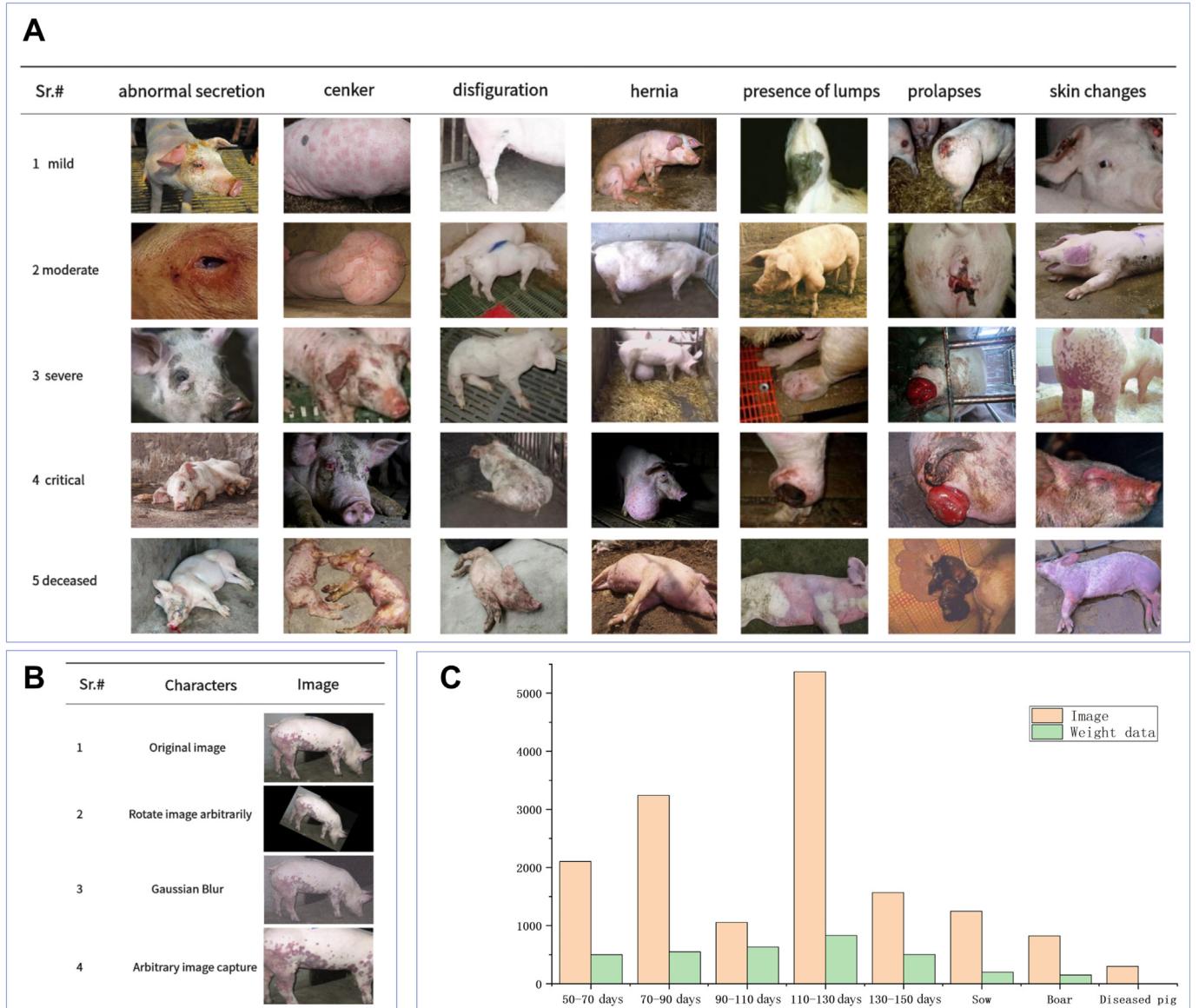


Fig. 5. (A) Different types and degrees of diseased pigs (images of sick pigs have been collected from the internet). (B) Image enhancement. (C) Data collection column chart.

were used for transfer learning on local datasets to facilitate the rapid training of the lower-level network. Additionally, techniques such as hyperparameter adjustment and freezing layers were adopted to improve the model performance; the application of the ResNet-50 network in this study is depicted in Fig. 7(B).

2.2.1. WGAN

Since their introduction by Goodfellow et al., 2020, the development of GANs has been rapid. Generative adversarial networks have been widely used in various fields, such as style transfer, image restoration, and font generation. Nash equilibrium theory is the core concept behind generating adversarial networks. In adversarial networks, both sides constantly confront each other to achieve their respective goals of maximizing interests. When the result of the confrontation reaches equilibrium, even if the optimization methods on both side are not necessarily globally optimal solutions, their results have practical significance. The architecture of GANs includes a generator and a discriminator network (Goodfellow et al., 2020). The generator and discriminator networks continuously compete with each other to reach the Nash equilibrium, as shown in Fig. 8

(A). Fig. 8(B) shows a generative network from a uniform distribution to a normal distribution during the learning process.

The generator is responsible for generating images that are close to the real samples. By learning the data distribution of the real sample, the generator generates a new sample $G(z)$ based on the acquired noise Z , which determines whether the input/output data belong to a real sample or a generated sample. When the discriminator receives image data x , the output $D(x)$ is determined as the probability that x is real. If $D(x) = 1$, x is a real sample, and if $D(x) = 0$, x is not a real sample. The GAN obtains a generative data distribution $P_G(x; \theta)$ as close as possible to the real data distribution $P_{\text{data}}(x)$ based on the random noise Z through a generator G . θ , which is a network parameter, adjusts the similarity between $P_{\text{data}}(x)$ and $P_G(x; \theta)$. Adversarial training between the generator and discriminator aims to minimize the respective loss values. The generator generates a data distribution that is as close to the real situation as possible and is agreed upon by the discriminator such that $D(G(z))$ tends to 1. At this point, the loss function value of the generator is minimized. The discriminator is expected to distinguish as accurately as possible whether the acquired data originate from the true or generated distribution, that is, $D(G(z))$ tends to zero and $D(x)$

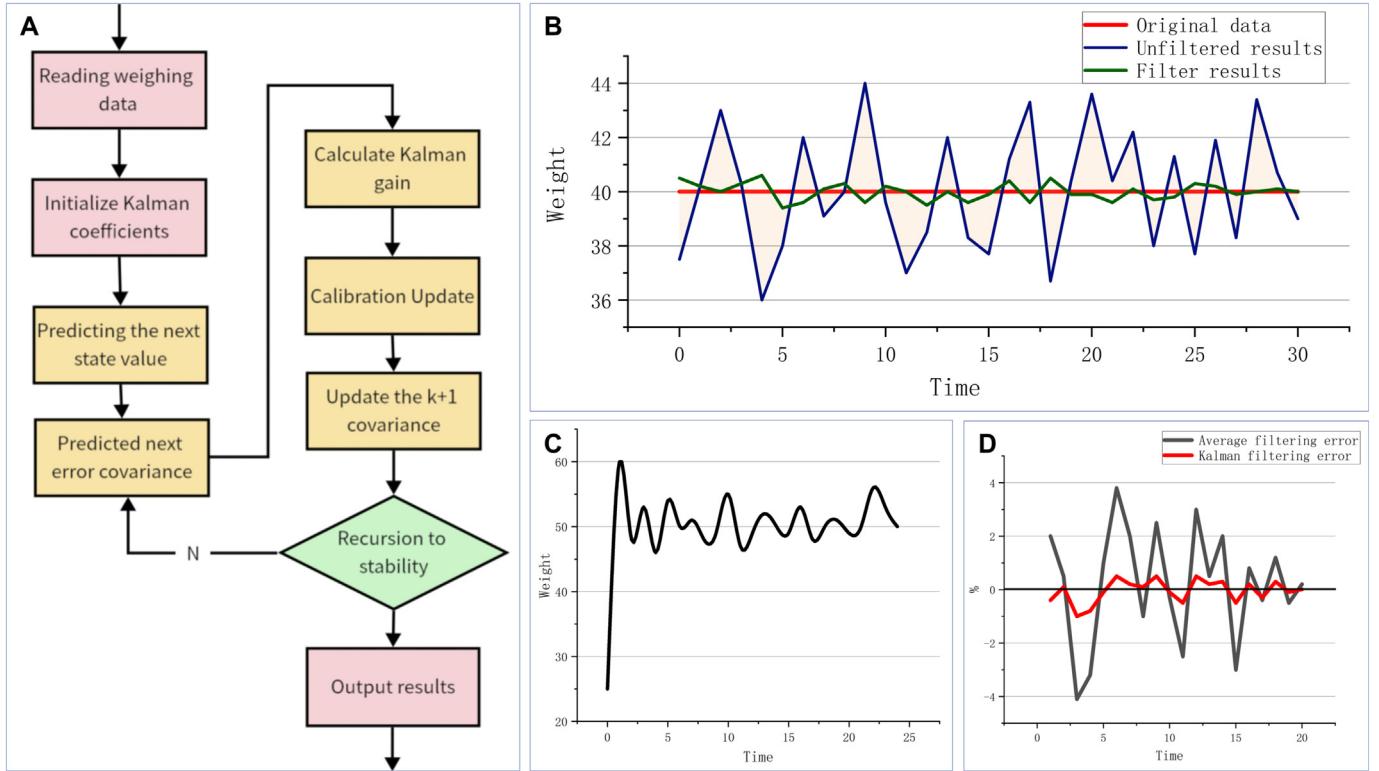


Fig. 6. (A) Kalman filtering process. (B) Comparison of Kalman filtering results when the true value is 40. (C) Actual weight waveform when the weight value is 50. (D) Comparison of average filtering error and Kalman filtering error for 20 sets of data.

tends to one. The loss functions of the generator and discriminator are constructed as follows:

$$\text{loss}_G = \log(1 - D(G(z))) \quad (6)$$

$$\text{loss}_D = -(\log D(x) + \log(1 - D(G(z)))) \quad (7)$$

Generator networks seek to generate optimal generators G_D^* as:

$$G_D^* = \arg \min_G \text{Div}(P_G, P_{\text{data}}), \quad (8)$$

The above equation indicates the scatter between the actual and generated data distributions, and the optimal generator refers to the minimum difference between the two distributions. In fixing the generator, the maximization objective function of the discriminator is expressed as:

$$\begin{aligned} V(G, D) &= E_{x \sim P_{\text{data}}(x)}[\log D(x)] + E_{z \sim P_G(z)}[\log(1 - D(G(z)))] \\ &= \int_x P_{\text{data}}(x) \log D(x) dx + \int_z P_G(x) \log(1 - D(x)) dx \\ &= \int_x [P_{\text{data}}(x) \log D(x) + P_G(x) \log(1 - D(x))] dx \end{aligned} \quad (9)$$

Discriminator networks seek to generate optimal discriminators D_G^* as:

$$D_G^* = \arg \max_D V(G, D) \quad (10)$$

Because the expected values of the generator and discriminator for $V(G, D)$ are exactly opposite, the generator and discriminator iterate continuously to achieve the optimal objective function of the overall network, solve the very small game, and achieve Nash equilibrium:

$$\min_G \max_D V(G, D) = E_{x \sim P_{\text{data}}(x)}[\log D(x)] + E_{z \sim P_G(z)}[\log(1 - D(G(z)))] \quad (11)$$

During the training process, the generator continuously improves its generative power to generate images that are as similar as possible to the real pattern to successfully fool the discriminator into thinking that it is the real image. The discriminator must continuously improve its discriminative power to avoid being fooled by the generator as much as possible. These two models iteratively play each other out to optimize the network and achieve a balance between them. When the generator is fixed, the discriminator is trained iteratively until it reaches the highest accuracy, achieving $D^*(x) = P_{\text{data}}(x)/(P_{\text{data}}(x) + P_G(x))$. After several iterations, the network model reaches the ideal state, where the distribution of the generated data is infinitely close to the actual data, achieving $P_{\text{data}}(x) = P_G(x)$, and the discriminator cannot recognize that the input data belong to the generator or dataset, thus achieving $D(G(z)) = 0.5$.

Although GANs have been greatly improved, they are characterized by easy collapse and difficult convergence owing to gradient disappearance and discriminator gradient invariance. Arjovsky et al. (2017) improved the GAN from the perspective of the loss function and proposed a novel Wasserstein distribution distance generative adversarial network (WGAN); they theoretically proved that the JS scatter (cross-entropy) interferes with the training stability of the original GAN. The JS scatter measures the distance between two probability distributions with non-overlapping regions and always obtains the constant $\log 2$, which causes the network gradient to disappear. The Wasserstein distance is formulated as follows:

$$W(P_r, P_g) = \inf_{\gamma \in \Pi(P_r, P_g)} E_{(x,y)}[\|x - y\|] \quad (12)$$

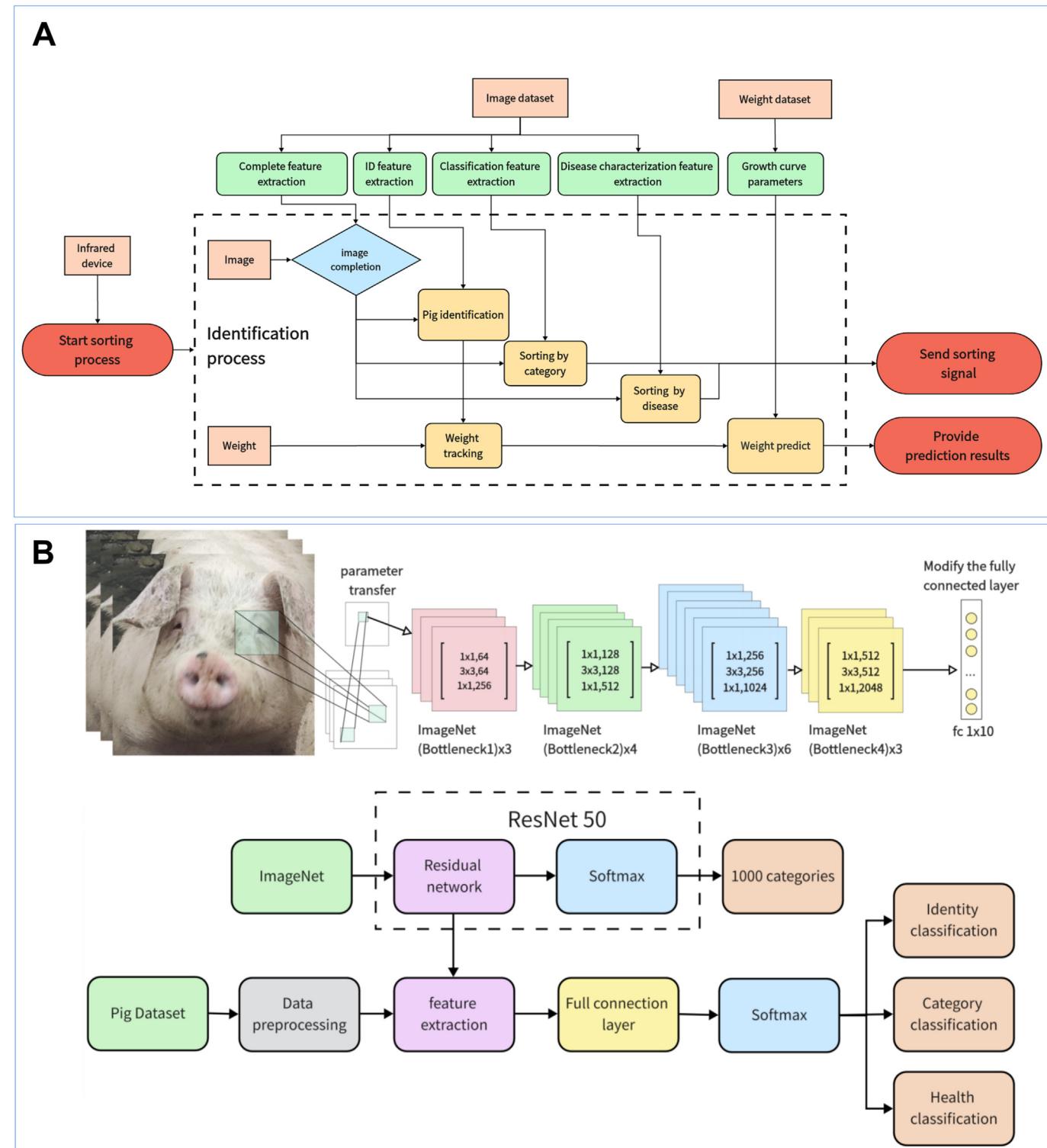


Fig. 7. (A) Overall flow chart of the feature extraction algorithm. (B) Network structure for livestock feature extraction using ResNet-50 as the backbone network.

where P_r is the real data distribution, P_g is the generated data distribution, and $\Pi(P_r, P_g)$ represents the set of joint data distributions formed by the combination of former two distributions; further, $(x, y) \sim \gamma$ is a distribution x from the data set and a distribution y from the generator obtained from the joint data distribution; $\|x - y\|$ is the spatial distance between distribution x and distribution y ; $E_{(x,y)}[\|x - y\|]$ represents the expectation value of the joint distribution for this spatial distance; $\inf_{\gamma \sim \Pi(P_r, P_g)} E_{(x,y)}[\|x - y\|]$ represents the lower bound of the joint

distribution on the ideal value of the distance. The Wasserstein distance allows the determination of the relationship between P_r and P_g without requiring an overlapping region. The use of the Wasserstein distance requires the Lipschitz continuity, and an additional limit is imposed on the continuity function such that it satisfies the Lipschitz continuity. Discriminator f_w is constructed with parameter w and a nonlinear activation function at the end of the output layer to restrict parameter w to a specific range and maximize the objective function.

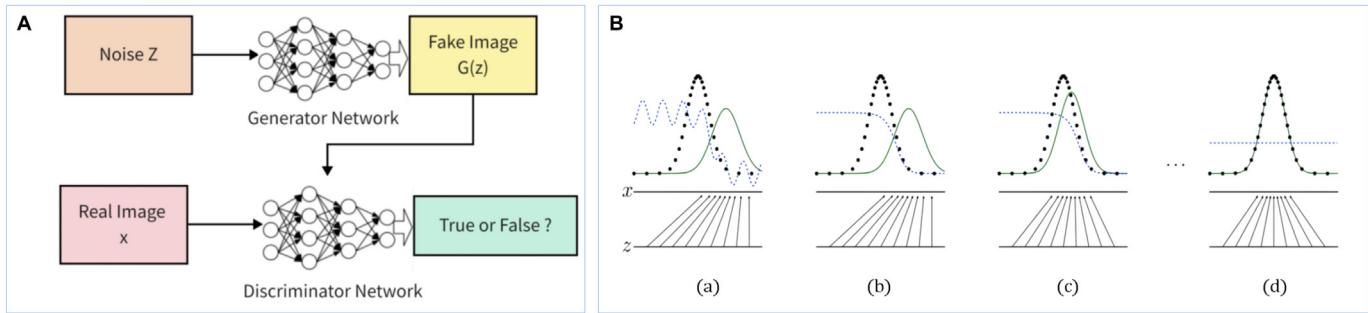


Fig. 8. (A) Generative adversarial network framework. (B) Generating a web-based learning process.

$$L = E_{x \sim P_r}[f_w(x)] - E_{x \sim P_g}[f_w(x)] \quad (13)$$

The generator loss (G loss) function L_G and discriminator loss (D loss) function L_D of the WGAN are defined as follows:

$$L_G = -E_{x \sim P_g}[f_w(x)] \quad (14)$$

$$L_D = E_{x \sim P_g}[f_w(x)] - E_{x \sim P_r}[f_w(x)] \quad (15)$$

Thus, the WGAN can solve the problems of the GAN, namely, training instability and collapse-mode diversity, thus ensuring diversity of the generated samples. In this study, a mask M was added to the input image as a broken image for the generator to generate the model, and the WGAN algorithm was used. Furthermore, the contextual loss function $L_{\text{contextual}}$ and the perceptual loss function $L_{\text{perceptual}}$ were added to ensure that the same input image was obtained, and the contextual and perceptual losses were combined with the argmin function as Z. Herein, λ functions as a hyperparameter and is used to control the degree of importance of the contextual loss compared to perceptual loss. In this study, the default setting is $\lambda = 0.1$. The losses are calculated as follows:

$$L_{\text{contextual}}(Z) = \|M \odot G(z) - M \odot y\| \quad (16)$$

$$L_{\text{perceptual}}(Z) = \log(1 - D(G(Z)))D(G(Z)) \quad (17)$$

$$L(z) = L_{\text{contextual}}(z) + \lambda L_{\text{perceptual}}(z) \quad (18)$$

$$z^* = \arg \min_z L(z) \quad (19)$$

In training, the weight of the GAN loss in the generator loss must be adjusted to ensure that the G loss and GAN loss are on equal footing, or failing that, the G loss is one scale larger than the GAN loss. At the same time, the ratio of the training time of the generator and to that of the discriminator is adjusted. Additionally, the discriminator is trained five times, while the generator is trained only once when conducting the experiment; the learning rate should not be excessively large, and training optimization is performed using the Adam optimizer.

2.2.2. Residual network model

He et al. (2016a, 2016b) proposed a ResNet network structure, also called a residual network, to solve the degradation problem of deep networks. The basic idea is to add the concept of residual learning to the traditional CNN to transform the learning of the output into the learning of the residuals to solve for the gradient dispersion and accuracy degradation in deep networks (He et al., 2016a, 2016b) and control the speed while ensuring accuracy (Krizhevsky et al., 2012). The residual network structure used is ResNet-50. The residual module first reduces 256 channels to 64 channels by 1×1 convolution and then recovers for the residual network unit by 1×1 convolution in 256 channels, as shown in Fig. 9(A). Fig. 9(B) shows a single residual network unit.

Deeper CNNs are used after applying the Visual Geometry Group (VGG) model. However, it has been experimentally shown that an increase in network depth causes the gradient to explode and disappear, resulting in a system that does not converge (Lee et al., 2019). In the VGG model, the network is guaranteed to converge by using stochastic gradient descent (SGD) in backpropagation through a data normalization operation on the input data and intermediate layers; however, when the depth of the model network exceeds a dozen layers, SGD becomes ineffectual. To address this problem, a deep residual network, ResNet, has been proposed, which allows the network to be as deep as possible. However, it feeds a portion of the input data directly to the output without going through the convolutional network, and retains some of the original information. The ResNet network can add deeper convolutional layers to enhance this effect, and the residual block structure plays a key role in its functioning.

If the dimension of the residual mapping $F(x)$ is different from that of the jump connection X , it cannot be summed up; therefore, X must be up-dimensioned to ensure equality of dimensions before summing up; thus, a 1×1 specific convolutional kernel with stride = 2 is usually added to the jump connection X to ensure that the output of X is the same as that of the convolutional block. When residual blocks are stacked repeatedly, network structures of different depths can be formed (Brito et al., 2019). After experimenting with various depths of networks, such as ResNet-18, ResNet-34, ResNet-50, ResNet-101, and ResNet-152, a 50-layer ResNet-50 network was selected for this study based on its practical effect and computational volume.

The principle behind the ResNet network is the assumption that a relatively shallow network has reached saturation accuracy and the addition of several constant mapping layers (identity mapping, $y = x$, output is equal to input) thereafter, increasing the network depth. However, the error does not increase; that is, a deeper network does not increase the error of the training set, and the residual structure is calculated as follows:

$$y_1 = h(x_1) + F(x_1, W_1), \quad (20)$$

$$x_{l+1} = f(y_l), \quad (21)$$

where x_1 and x_{l+1} denote the input and output of the first residual unit, respectively, where each residual unit contains a multilayer structure; F is the residual function, which denotes the learned residual; and $H(x) = X_1$ denotes the constant mapping; and f is the ReLU activation function. The learned features from shallow layer 1 to deep layer L can be obtained as:

$$x_L = x_t + \sum_{i=1}^{L-1} F(x_i, W_i) \quad (22)$$

Using the chain rule, the gradient of the inverse process can be obtained as:

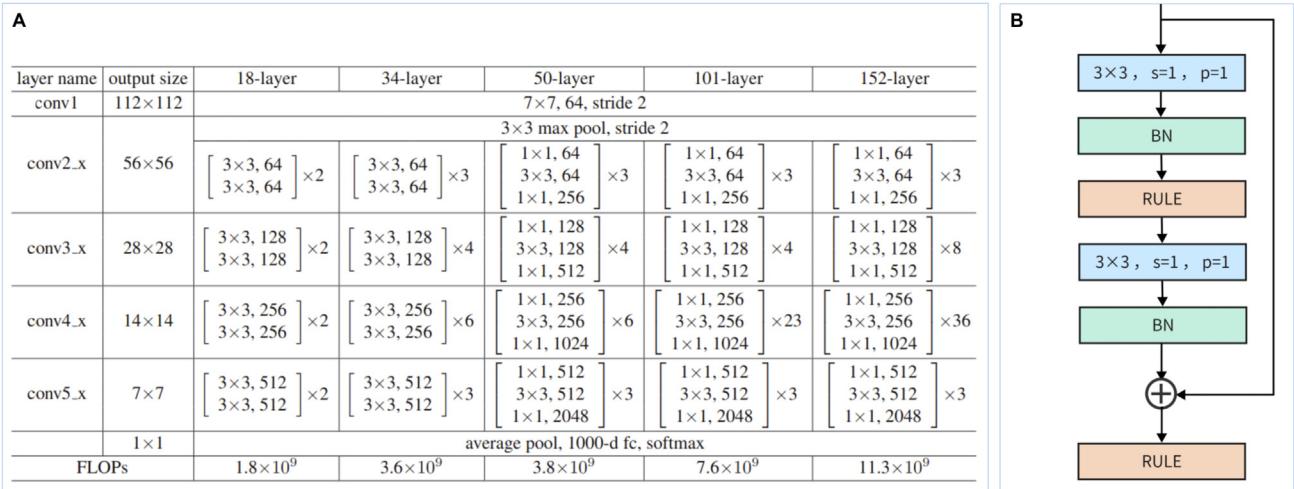


Fig. 9. (A) ResNet network. (B) Residual network unit.

$$\frac{\partial \text{loss}}{\partial x_i} = \frac{\partial \text{loss}}{\partial x_L} \cdot \frac{\partial x_L}{\partial x_i} = \frac{\partial \text{loss}}{\partial x_L} \cdot \left(1 + \frac{\partial \text{loss}^L}{\partial x_L} \sum_{i=1}^{L-1} F(x_i, W_i) \right) \quad (23)$$

The first factor $\frac{\partial \text{loss}}{\partial x_L}$ in the above equation indicates the gradient reached by the loss function, and the “1” in parentheses indicates that the short-circuiting mechanism can propagate the gradient without loss, while the other residual gradient must pass through the layer with weights, and the gradient is not passed directly. The residual gradient is incomplete. Even if the value is small, the presence of 1 does not cause the gradient to vanish, making residual learning easier.

The main structure of the ResNet-50 network consists of one convolutional layer and four residual blocks, each containing several bottleneck residual units with different numbers of frozen residual unit parameters.

2.2.3. Model training

In the identity classification experiment, pig identification was performed using the ResNet-50-based model pre-trained on the ImageNet dataset. After pre-training on the ImageNet dataset, the convolutional layer of the ResNet-50 model was frozen, and the parameters of the fully connected layer were retrained to obtain the accuracy of the network. The captured images were classified and saved for filtering and cropping, with 6000 images as the training set, 2000 images as the validation set, and 2000 images as the test dataset.

In the category classification experiment, the pigs were classified into five categories according to their growth and development. Five groups of classified images, each containing 1800 images, were used to divide the training, validation, and test sets in a 3:1:1 ratio. The initial learning rate used for the experimental training of all CNNs was 0.003, and the loss function was a cross-entropy loss function.

This study expanded the dataset to 1200 images by screening 300 images of sick pigs. The dataset was expanded to 800 images using arbitrary flipping, Gaussian noise, and cropping and mixed with 1200 images of healthy pigs to divide the training, validation, and test sets in a ratio of 3:1:1. The pre-trained models were migrated to the target network for training. Most sick pigs develop lesions on their face and back, which can be sorted out to isolate the sick pigs through specific fences. The camera unit of the sorting system captures images of the back and face in a model trained after pre-training and fine-tuning. The two images detected simultaneously were assigned to the prediction model, and the results obtained were fused to obtain the classification results. The weights of the detection results for the face and body

images are 40% and 60%, respectively. The fused results were obtained as the detection results of the diseased pigs.

2.3. Data prediction

In this study, age and body weight data were obtained by aggregating data collected from field and public datasets. The main reference for prediction by fitting growth curves is the less disturbed and most vigorous developmental age of 50–150 days. Fig. 10(A) shows a scatter plot of the weight data used in this study, and Fig. 10(B) shows the cumulative weight and growth coefficient of the pigs counted once every 10 days.

Growth curves for the body weight of the pigs were obtained and fitted nonlinearly using a logistic model with the following equation:

$$W = A / (1 + e^{(a - rt)}) \quad (24)$$

$$R^2 = 1 - \sum (y - \hat{y})^2 / \sum (y - \bar{y})^2 \quad (25)$$

Here, W is the weight at t days of age (kg), A is the limiting weight parameter (kg), and b is a constant scale, r is the growth rate parameter, and e is a constant. The pig coefficient parameter equation is expressed as follows:

$$c = \frac{W_{t2} - W_{t1}}{W_{t2} \times \Delta t} \quad (26)$$

The model calculates the following parameters in logistic regression: K = 154.60; a = 26.75; r = 0.022; R² = 0.9956. The pigs were divided into six categories: low development level (C1), sub-low development level (C2), normal development level (C3), higher development level (C4), high development level (C5), and abnormal development (C6). The absolute weight and recent daily growth coefficients obtained from the last two statistics are examined in the study, both of which are below 90% of the normal levels for low development level; one is below 90% of the normal level and the other is below average for the sub-low developmental level; both are above 90% of the normal levels for the high development level, and one is above 90% of the normal level and the other is above average for the high developmental level. The rest were classified as having a normal development level, as well as abnormal development owing to factors such as disease. Through the classification of five categories characterizing the growth level of the pigs, a statistical analysis of the integration of the five categories of

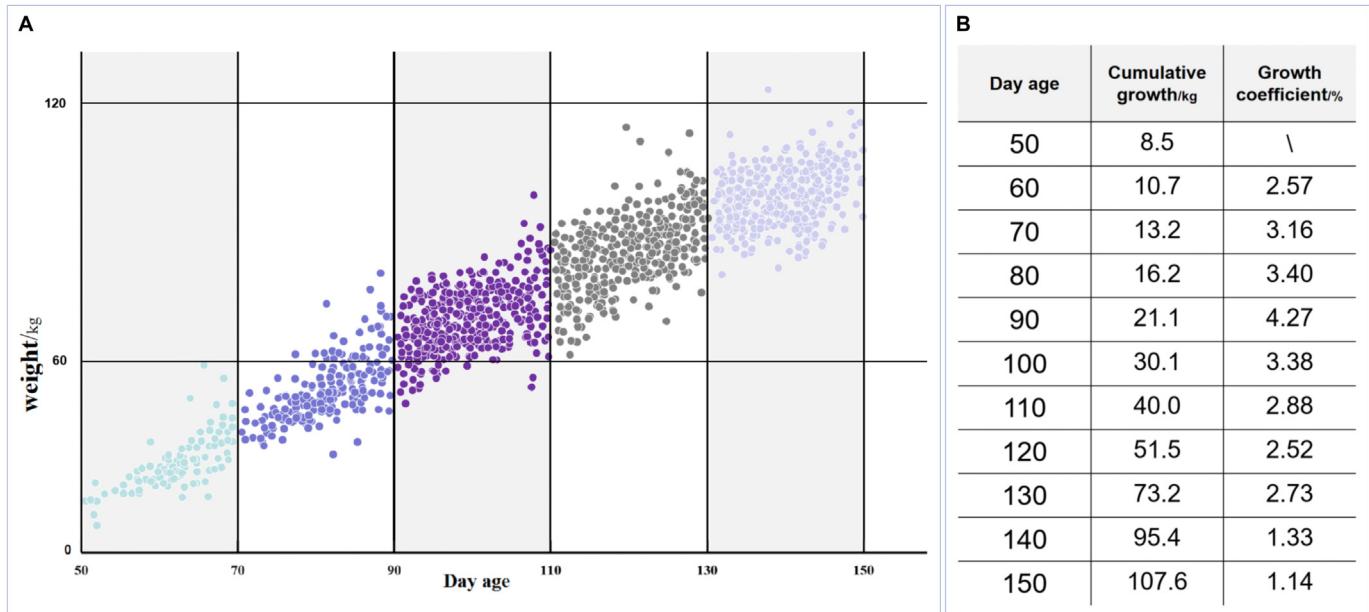


Fig. 10. (A) Pig weight scatter plot. (B) Pig weight data.

normal development and abnormal onset was performed. In the present study, a five-category sorting mechanism was designed for pigs at different developmental stages to evaluate and predict their growth using weight and growth coefficients.

By fitting the growth curves, scientific statistics and predictions were obtained for the studied pigs. A folded histogram of the statistics is shown in Fig. 11(A). The bar chart shows the cumulative growth as an S-shaped curve, where the line graph denotes the growth coefficient,

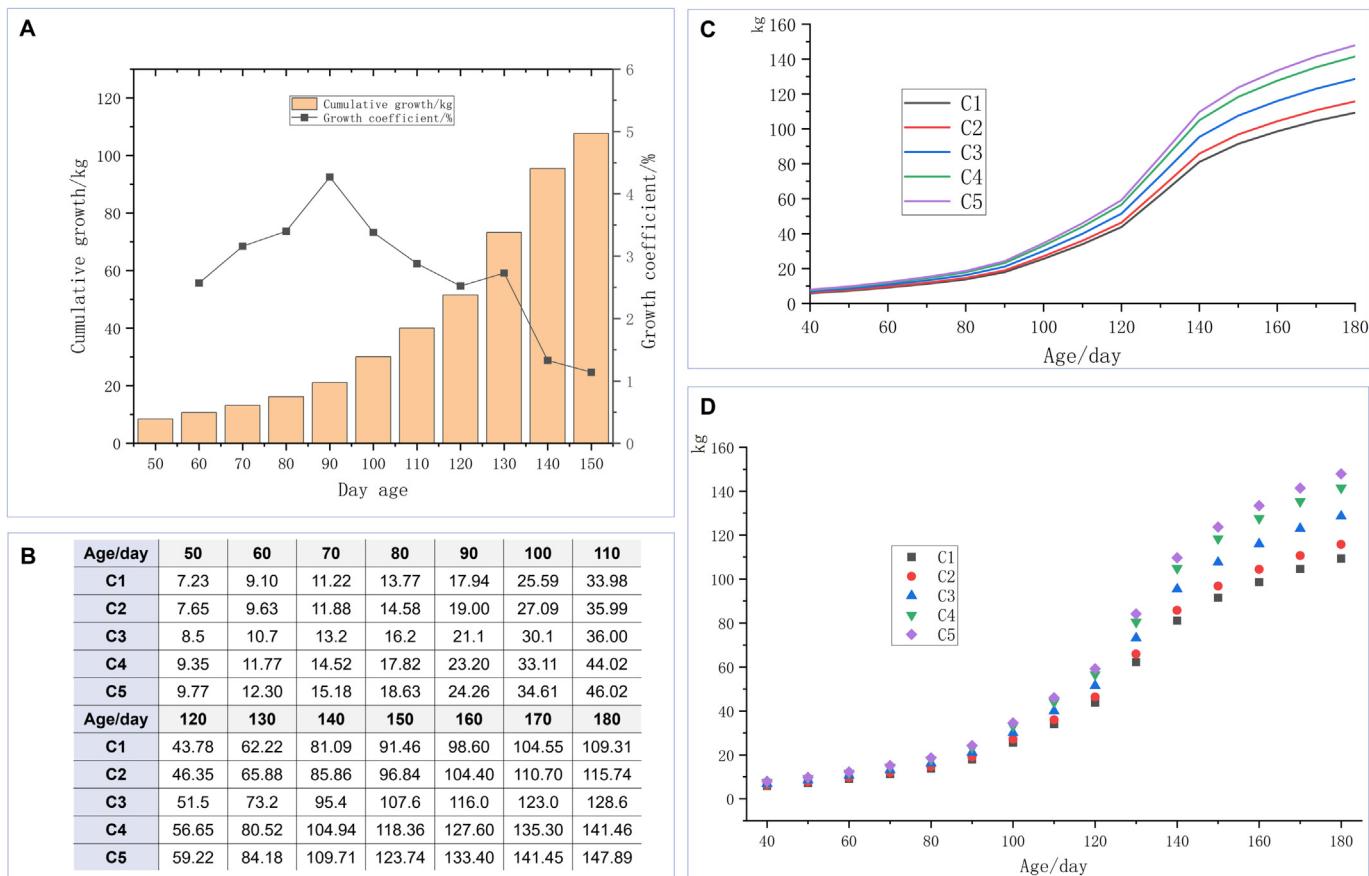


Fig. 11. (A) Cumulative growth and growth coefficient of pigs from 50 to 150 days of age. (B) Predicted weight data (kg) of pigs in different classes at each stage after classification based on five categories. (C) Line graph of predicted weight data for each class. (D) Point plots of predicted weight data for each class.

which first increases and then decreases. The weights of the pigs from 50 to 180 days of age are also predicted based on the fitted growth curve, and the data are shown in Fig. 11(B). The line and points plots for the predicted weight are shown in Figs. 11(C) and (D), respectively.

Five categories of normal pigs and one category of abnormal pigs were classified, and the five statistical categories of normal development and abnormal morbidity were integrated and analyzed. Numerous pieces of information regarding pigs can be obtained using the above method, which are aggregated to design a sorting monitoring system.

A visualization interface is shown in Fig. 12. The visualization interface can display livestock data more intuitively and automate sorting management.

3. Results and discussion

3.1. Image training results

Image recognition comprises three components: individual recognition, classification recognition, and diseases recognition, all with good recognition accuracies. For the effect of occlusion on image recognition, the WGAN image complementation algorithm is used, and the feasibility of the experimental design is demonstrated by comparing the source, occlusion data, and complementation data.

3.1.1. Identification results

The images in the test dataset were fed into the trained pig identification model for pig identification during testing, and the images successfully identified were deposited into the corresponding dataset. The highest accuracy was obtained in the first 25 frozen layers. Simultaneously, it is necessary to make a timely update the dataset. As the pigs grow and develop rapidly, increasing the sorting frequency guarantees identity recognition accuracy. The training was performed for 100 epochs, 750 iterations, and a batch size of 8. The learning rate was 0.005, and ReLU was used as the activation function. The loss function versus accuracy curves for the training and test sets are shown in Figs. 13(A) and (B), respectively.

Identity recognition is highly dependent on the environment, and the recognition accuracy drops drastically when the image is occluded.

The image complementation method of WGAN was used to add a mask to the original input image, combining context and perceptual losses. Context loss ensures that the obtained image is identical to the input image, and perceptual loss ensures the authenticity of the image used to find the complementary restored image. In this study, the pre-trained WGAN model was first tested on the LSUN dataset for image reconstruction, based on which, the pre-trained model was used to complete the image restoration task. The effect of image restoration is shown in Fig. 14(A).

In the experiment, the complementary images were retested, and a significant improvement in the accuracy rate was observed, as shown in Fig. 14(B). The overall recognition rate of the occluded data is only 60.3%, whereas the recognition rate of the complemented image reaches 82.3%, which is close to the recognition rate of 88.9% for the original image, thus satisfying the requirements of the experimental design and improving the robustness of the classification model. A sample of ten identified pigs was selected to plot the confusion matrix of the original and repaired data, as shown in Figs. 15(A) and (B), respectively.

3.1.2. Sorting results and disease recognition

In the experiment, the network model was trained according to the pig classification requirements. The accuracy of ResNet-50 was verified by freezing 3, 7, and 13 layers; the experimental results are shown in Fig. 16(A). The results demonstrate that the highest training results are obtained at the 91st cycle for 7 frozen layers. The average accuracy of the experimental models exceed 95%, and optimized ResNet-50 was used to achieve accurate predictions for different classifications to meet the practical needs of sorting. To show the classification effect more intuitively, tSNE values for D1 to D5 are plotted based on the time series in Figs. 16(C) and (D) and tSNE values for C1 to C5 are plotted based on the level series in Figs. 16(E) and (F).

The model was retrained on a disease recognition dataset, the underlying network was frozen, and the training was completed and tested on a test set. The accuracy of the disease test set is 97.81%, which satisfies the experimental requirements, as shown in Fig. 16(B). The confusion matrices for the training and test set predictions are also plotted in Figs. 17 (A) and (B), respectively.

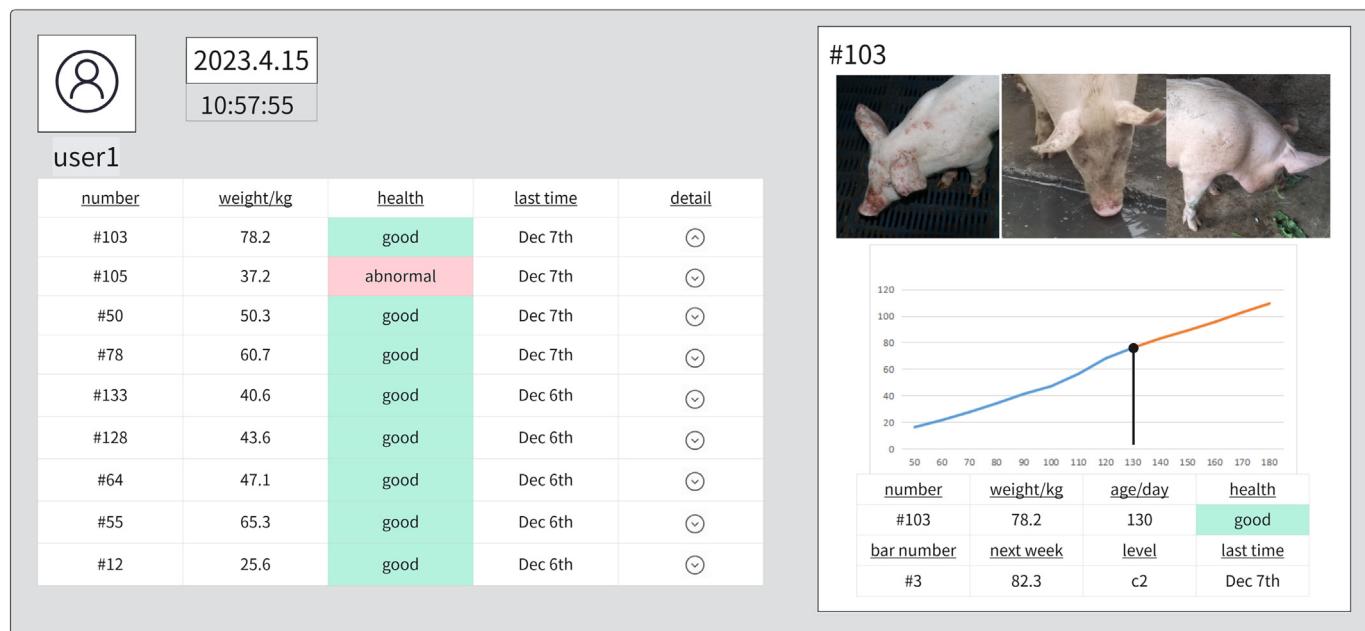


Fig. 12. Designing visual interfaces.

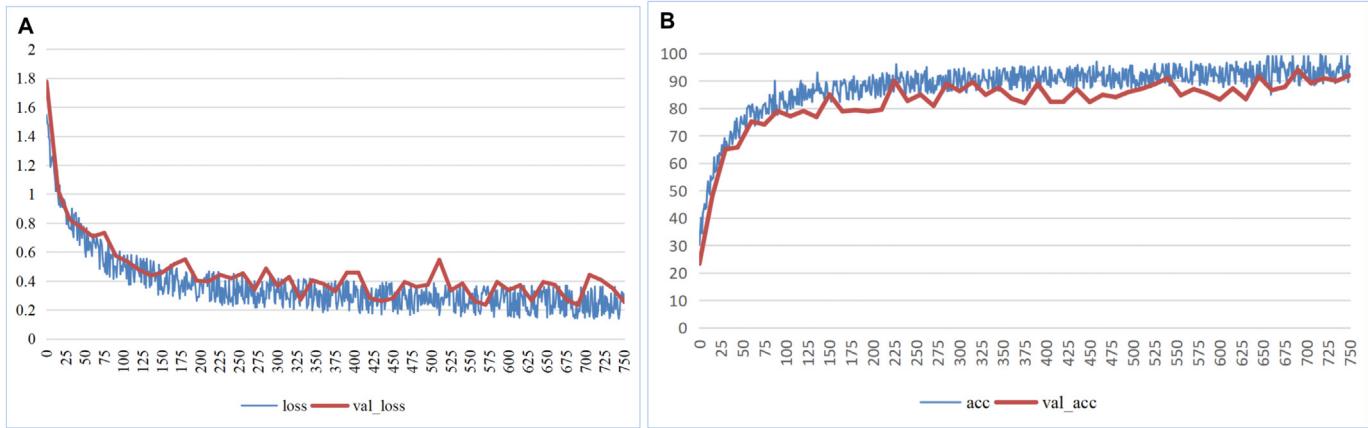


Fig. 13. (A) Loss value curve for identity model training. (B) Accuracy of model training.

3.2. Experimental results

Experimental results are obtained based on fine-tuning transfer learning using a residual CNN model for classification and sorting. The accuracy of the pig ID recognition reaches 89%. The proposed system collects pictures of different types of pig diseases and learns binary classification from healthy pigs. Considering a specific situation in which the image is incomplete owing to stray light and railings in the application scenario, the WGAN algorithm is used to improve the image quality and sorting accuracy. The ResNet-50 network is used for individual identification. Transfer learning can quickly learn and deploy models with a limited sample size, effectively improving system effectiveness. To fit the growth curves of pigs, the absolute weight and growth between two detections are considered, the pigs are categorized (six categories) based on the growth and development stages, and the pig growth trends are evaluated for farm management reference. The experimental results show that the classification success rate for nursery pigs, piglets, and fattening pigs reaches 95%, whereas the classification success rate for sick and healthy pigs reaches 97%. Using image completion algorithms to supplement occluded images can significantly improve accuracy by approximately 32% (Table 3).

To better evaluate the effects of model identification, the accuracy, precision, sensitivity, F1 value, and consistency test parameters were calculated for the four sorting types, as listed in Table 2. The parameters of the model identification results are consistent with the experimental results.

Following equations describe the calculation of the test parameters:

$$\text{Precision} = 100 \times \left(\frac{T_p}{T_p + F_p} \right) \quad (27)$$

$$\text{Sensitivity} = 100 \times \left(\frac{T_p}{T_p + F_N} \right) \quad (28)$$

$$F1 - \text{Score} = 100 \times \left(2 \times \left(\frac{\text{Precision} \times \text{Sensitivity}}{\text{Precision} + \text{Sensitivity}} \right) \right) \quad (29)$$

$$\text{Accuracy} = 100 \times \left(\frac{T_p + T_N}{T_p + T_N + F_p + F_N} \right) \quad (30)$$

Kappa is a statistical coefficient used in evaluation measures to validate results. The kappa values illustrate the consistency of the predicted values with the actual values. The kappa values remain between zero and one, indicating the confidence level of the given rule. Lower kappa values indicate no agreement; however, the low agreement obtained in the experiment in this study is informative and within a reasonable interval. Kappa is calculated as follows:

$$\kappa = \frac{p_0 - p_e}{1 - p_e}, \quad (31)$$

where p_0 is the sum of the number of correctly classified samples in each category divided by the total number of samples, that is, the overall classification accuracy, C is the total number of categories, and T_i is the

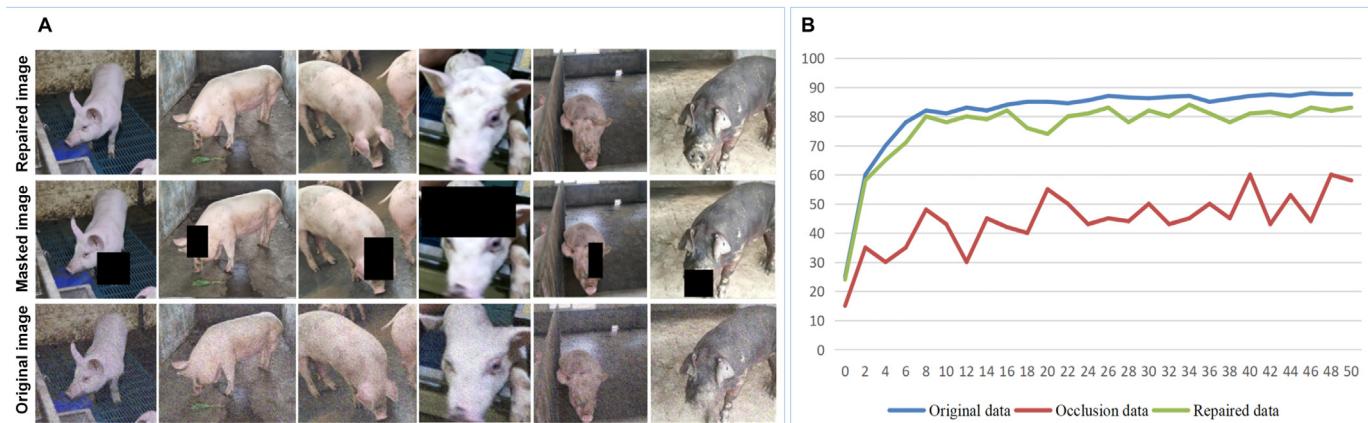


Fig. 14. (A) Comparison of image complementation results. (B) Three types of image recognition results.

		Predicted Class									
		E1	E2	E3	E4	E5	E6	E7	E8	E9	E10
True Class	E1	93.5	0.0	1.2	1.1	0.8	0.0	1.8	0.0	1.6	0.0
	E2	0.0	89.1	0.0	0.0	1.9	0.0	1.7	2.2	0.0	5.1
	E3	1.1	0.0	70.2	0.0	0.0	3.3	0.0	2.4	0.0	23.0
	E4	1.6	1.3	0.0	91.3	0.0	0.0	1.9	0.0	10.7	3.2
	E5	0.0	1.4	2.0	0.0	89.4	0.0	0.0	4.0	0.0	3.2
	E6	3.9	0.0	3.1	0.0	0.0	91.2	0.0	0.0	1.0	0.8
	E7	1.9	1.5	0.0	1.7	0.0	0.0	88.4	0.0	2.0	4.5
	E8	0.0	2.0	2.4	0.0	3.9	0.0	0.0	89.3	0.0	2.4
	E9	3.1	0.0	0.0	2.3	0.0	1.1	1.8	0.0	90.6	1.1
	E10	0.0	0.5	0.0	2.0	0.2	1.3	0.0	0.0	0.0	96.0

		Predicted Class									
		E1	E2	E3	E4	E5	E6	E7	E8	E9	E10
True Class	E1	83.6	0.0	1.7	1.9	3.2	0.0	2.8	0.0	3.0	3.8
	E2	0.0	84.4	0.0	1.3	1.6	0.0	2.1	2.2	0.0	8.4
	E3	1.3	0.0	84.5	0.0	0.0	1.2	0.0	3.0	0.0	10.0
	E4	1.7	1.4	0.0	82.1	0.0	0.0	2.2	0.0	4.6	8.0
	E5	0.0	1.3	2.0	0.0	83.0	0.0	0.0	5.0	0.0	8.7
	E6	3.9	0.0	3.1	0.0	0.0	83.1	0.0	0.0	1.8	8.1
	E7	2.0	1.6	0.0	2.0	0.0	0.0	84.6	0.0	2.2	7.6
	E8	0.0	2.1	3.0	0.0	4.9	0.0	0.0	82.4	0.0	7.6
	E9	3.1	0.0	0.0	9.3	0.0	1.9	1.8	0.0	75.6	8.3
	E10	0.0	0.7	0.0	1.2	0.0	8.0	0.0	0.0	0.0	90.1

Fig. 15. (A) Confusion matrix for identification of original data. (B) Confusion matrix for identification of repaired data.

number of correctly classified samples in each category. Supposing that the number of true samples in each category is a_1, a_2, \dots, a_C , the number of predicted samples in each category is b_1, b_2, \dots, b_C , and the total number of samples is n , then:

$$p_o = \frac{\sum_{i=1}^C T_i}{n} \quad (32)$$

$$p_e = \frac{\sum_{i=1}^C a_i \times b_i}{n^2} \quad (33)$$

Compared with other models, the livestock identification model based on the residual and generative adversarial networks designed in this study has better features and more advantages. In the field of pig recognition, the recognition rate obtained in this study for covered images is significantly higher than that obtained in previous studies and close to the recognition accuracy of the original images. Additionally a good classification effect in terms of classification, sorting, and characterization of disease detection is obtained in this study.

Based on the above results, it can be concluded that the livestock identification and information management system proposed in this study provides satisfactory results in terms pig identification and

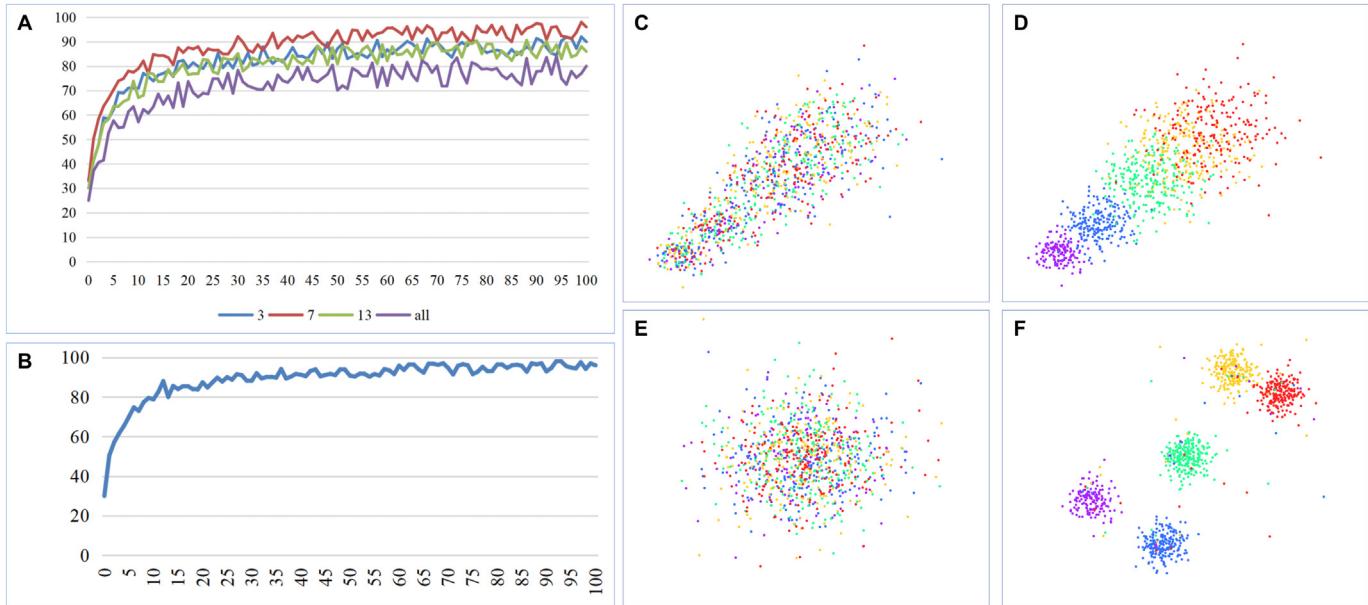


Fig. 16. (A) Classification results under different freezing layers. (B) Training curve for disease recognition. (C) Confused tSNE chart (time series) for five types of pigs. (D) tSNE chart (time series) for five types of pigs after classification. (E) Confused tSNE chart (level series) for five types of pigs. (F) tSNE chart (level series) for five types of pigs after classification.

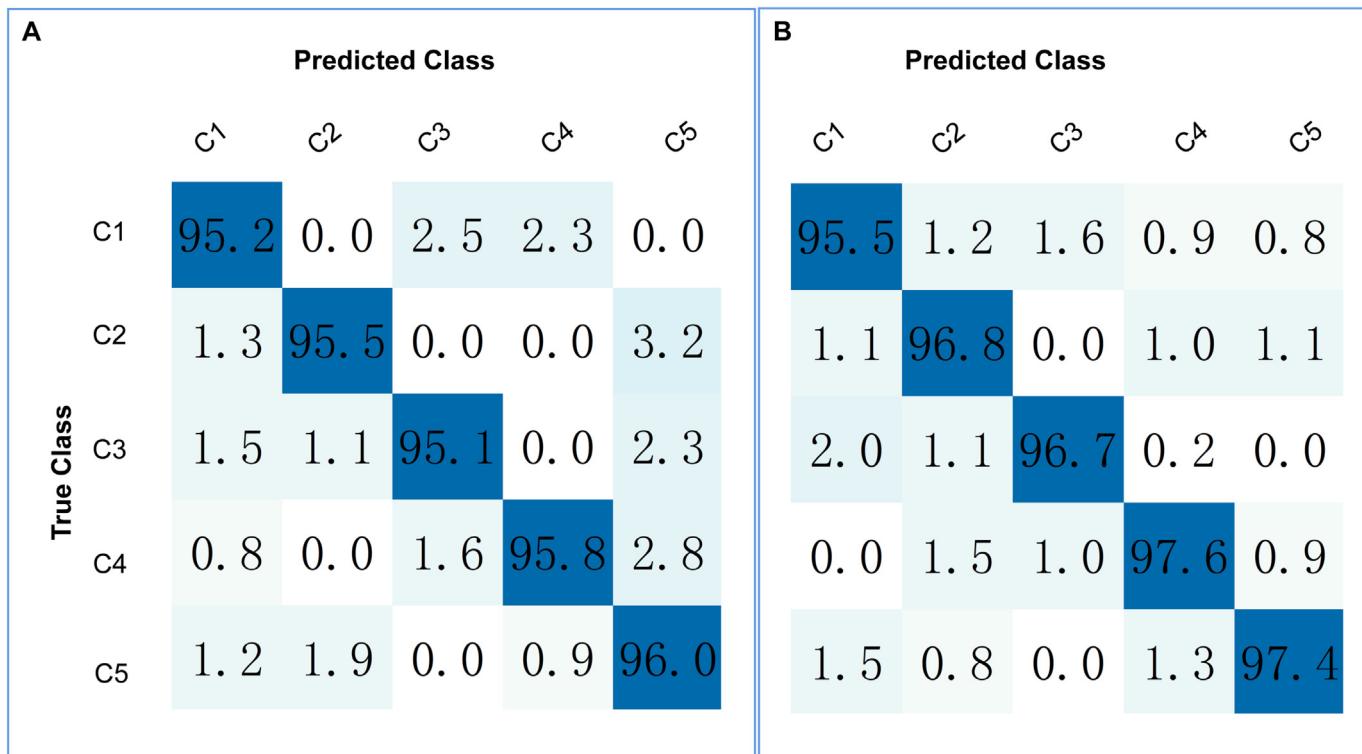


Fig. 17. (A) Confusion matrix for five-class classification recognition training set. (B) Confusion matrix for five-class classification recognition test set.

information processing compared with those of similar systems, especially with respect to occluded image recognition. The proposed deep-learning network model exhibits good results and may have significant implication in the field of in livestock sorting and information management.

4. Conclusion

In this study, a low-cost livestock-sorting system for image recognition was designed based on deep learning. Using only weights and images as inputs, the proposed system accomplishes functions such as livestock ID recognition, automated livestock sorting, disease detection, and data tracking and prediction, through the technical route of deep learning and the industrial Internet of Things. The following conclusions are drawn from this study.

(1) A deep-learning-based machine vision sorting system is proposed and successfully tested for processing field images with multiple feature recognition tasks, such as disease and illness, sex and reproduction, growth, and fattening. It uses deep learning to complete the individual recognition of livestock, performs

category sorting, and detects diseases based on image data. It can be fully applied to all aspects of sorting including improving sorting efficiency, reducing the workload of manual sorting, and preventing bacterial infection.

- (2) Deep-learning technology is used to collect images and build different datasets for different requirements. The ResNet-50 model is used to ensure the accuracy and generalization of the results through pre-processing models and parameter fine-tuning. Different models are trained for different requirements to meet the system requirements.
- (3) GANs are applied to the ResNet backbone model to improve the recognition accuracy against occluded images, allowing the model to be applied to a wider range of scenarios.
- (4) A weight health evaluation system is designed that encompasses modules such as data collection, individual recognition, accuracy verification, health and development evaluation, and weight prediction. The essential functions of the system are then completed.
- (5) A Kalman filter-based filtering algorithm is developed to improve the accuracy of the dynamic process of the deep-

Table 2
Result evaluation parameter.

Studies	Year	Species	Objects	Backbone	Accuracy
Andrew et al.	2021	cow	Individual recognition	CNN	93.80%
Pan et al.	2022	cow	Individual recognition	CNN	93% (recent variants 85%)
Xu et al.	2022	cow	Individual recognition	CattleFaceNet	91.35%
Wang et al.	2022	pig	Individual recognition	DenseNet 121	94.04%
Marsot et al.	2020	pig	Individual recognition	CNN	83.70%
Wang et al.	2021	pig	Individual recognition	ResNet	67.58% (covered)
			Individual recognition	ResNet+WGAN	82.9% (covered)
			Individual recognition	ResNet+WGAN	88.90%
Ours current study	2023	pig	Classification recognition	ResNet+WGAN	95%
			Disease recognition	ResNet+WGAN	98%

Table 3

Comparison of livestock identification under different models.

Method	Accuracy(%)	Precision(%)	Sensitivity(%)	F1-Score(%)	Kappa
ID-Original	88.92	88.75	88.89	88.50	0.62
ID-Repaired	83.34	83.26	83.49	83.27	0.60
Classification	95.20	94.92	95.12	94.88	0.65
Health	97.81	97.89	97.66	97.56	0.67

learning-based sorting system for multi-purpose applications on pig farms.

However, the generalizability of the model to different livestock is worth further investigation because of the significant cost of data collection and tagging, as well as the need for a more robust and transferable model. More extensive data must be collected to train and validate the reliability of the model. The potential of heavy recognition and 3D recognition technologies in the field of livestock sorting based on image recognition has been recognized, which may enable the modeling of livestock body types and improve recognition accuracy. Because of the lack of collaboration in the veterinary field at this stage, this study only focuses on the classification and detection of diseases with obvious livestock phenotypes; diseases related to livestock pathology and multi-information fusion detection should be the focus of future research. Furthermore, a collaboration with the Institute of Animal Husbandry and Veterinary Science on a cross-cutting project targeting livestock pathogenesis, processes, and image recognition is anticipated. Moreover, another collaboration with Ningxia Jinyuhaoxing Agriculture and Animal Husbandry Co. Ltd. and Ningxia Yanchi Tan Sheep Industry Development Group Co. Ltd. is in the pipeline focusing on collecting data on other livestock, including dairy cows and tan sheep, to realize research and innovation in sorting information management systems for other livestock breeds.

Funding

This work was financial supported by the Zhenjiang Science and Technology Bureau, Jiangsu Science and Technology Department, financial support was received from Zhenjiang Hongxiang Automation Technology Co. Ltd.

Institutional review board statement

Not applicable.

Informed consent statement

Not applicable.

CRediT authorship contribution statement

Yuanzhi Pan: Conceptualization, Data curation, Formal analysis, Funding acquisition, Methodology, Resources, Software, Validation, Visualization, Writing – original draft, Writing – review & editing. **Yuzhen Zhang:** Data curation, Formal analysis, Methodology, Software, Validation, Visualization, Writing – original draft. **Xiaoping Wang:** Project administration. **Xiang Xiang Gao:** Supervision. **Zhongyu Hou:** Writing – review & editing.

Data availability statement

Dataset is available at <https://data.mendeley.com/datasets/vd5vmgr8kg> (January 18, 2023) <https://data.mendeley.com/datasets/jy6hngx7df> (April 5, 2023).

Declaration of Competing Interest

The authors declare no conflicts of interest.

References

- Arjovsky, M., Chintala, S., Bottou, L., 2017. Wasserstein generative adversarial networks [C]/international conference on machine learning. PMLR 214–223.
- Brito, C., Machado, A., Sousa, A., 2019. Electrocardiogram beat-classification based on a ResNet network. *Stud. Health Technol. Informatics* 30 (6), 264–265.
- Calderón Díaz, J.A., García Manzanilla, E., Diana, A., et al., 2018. Cross-fostering implications for pig mortality, welfare and performance[J]. *Front. Vet. Sci.* 5, 123.
- Dong, C., Loy, C.C., He, K., et al., 2015. Image super-resolution using deep convolutional networks. *IEEE Trans. Pattern Anal. Mach. Intell.* 38 (2), 295–307.
- Fuentes, S., Viejo, C.G., Tongson, E., Dunshea, F.R., 2022. The livestock farming digital transformation: implementation of new and emerging technologies using artificial intelligence. *Anim. Health Res. Rev.* 1–13.
- Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., et al., 2020. Generative adversarial networks. *Commun. ACM* 63 (11), 139–144.
- He, K., Zhang, X., Ren, S., et al., 2016a. Deep Residual Learning for Image Recognition. IEEE.
- He, K., Zhang, X., Ren, S., Sun, J., 2016b. Deep residual learning for image recognition. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, USA.
- He, K., Gkioxari, G., Dollar, P., et al., 2017. Mask R-CNN. *Proceedings of the IEEE international conference on computer vision*, pp. 2961–2969.
- Kashiha, M., Bahr, C., Ott, S., et al., 2013. Automatic identification of marked pigs in a pen using image pattern recognition[J]. *Comput. Electron. Agric.* 93, 111–120.
- Kristensen, A.R., Nielsen, L., Nielsen, M.S., 2012. Optimal slaughter pig marketing with emphasis on information from on-line live weight assessment. *Livest. Sci.* 145 (1–3), 95–108.
- Krizhevsky, A., Sutskever, I., Hinton, G., 2012. ImageNet Classification with Deep Convolutional Neural Networks. NIPS.
- Kwon, Y.H., Casebolt, J.B., 2006. Effects of light refraction on the accuracy of camera calibration and reconstruction in underwater motion analysis. *Sports Biomechanics* 5 (2), 315–340.
- Latino, L.R., Pica-Ciamarra, U., Wisser, D., 2020. Africa: the livestock revolution urbanizes. *Global Food Security* 26.
- Lee, J.H., Kim, Y.J., Kim, Y.W., et al., 2019. Spotting malignancies from gastric endoscopic images using deep learning. *Surg. Endosc.* 89 (4), 806–815.
- Marsot, M., Mei, J.Q., Shan, X.C., et al., 2020. An adaptive pig face recognition approach using convolutional neural networks. *Comput. Electron. Agric.* 173.
- Miller, E.G., Matsakis, N.E., Viola, P.A., 2000. Learning from one example through shared densities on transforms. *Proceedings IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 464–471 (Hilton Head).
- Moekti, G.R., 2020. Industrial livestock production: A review on advantages and disadvantages. *IOP Conference Series: Earth and Environmental Science*. vol. 492. IOP Publishing, p. 012094 No. 1.
- Neethirajan, S., 2020. The role of sensors, big data and machine learning in modern animal farming. *Sensing Bio Sens. Res.* 29, 100367.
- Neethirajan, S., 2022. Affective state recognition in livestock—artificial intelligence approaches. *Animals* 12 (6), 759.
- Ouyang, H., Ren, L., 2023. Special issue “state-of-the-art porcine virus research in China”. *Viruses* 15 (2), 412.
- Pan, S.J., Tsang, I.W., Kwok, J.T., et al., 2011. Domain adaptation via transfer component analysis. *IEEE Trans. Neural Netw.* 22 (2), 199–210.
- Pan, Y., Jin, H., Gao, J., Rauf, H.T., 2022. Identification of Buffalo breeds using self-activated-based improved convolutional neural networks. *Agriculture* 12 (9), 1386.
- Pandey, S., Kalwa, U., Kong, T., Guo, B., Gauger, P.C., Peters, D.J., Yoon, K.J., 2021. Behavioral monitoring tool for pig farmers: ear tag sensors, machine intelligence, and technology adoption roadmap. *Animals* 11 (9).
- Radford, A., Metz, L., Chintala, S., 2015. Unsupervised representation learning with deep convolutional generative adversarial networks. *arXiv preprint arXiv:1511.06434*.
- Robbins, R.C., Almond, G., Byers, E., 2014. Swine diseases and disorders[J]. *Encyclopedia Agricult. Food Syst.* 261.
- Tzeng, E., Hoffman, J., Darrell, T., et al., 2015. Simultaneous deep transfer across domains and tasks. *Proceedings 2015 IEEE International Conference on Computer Vision (ICCV)*. IEEE, Santiago, Chile, pp. 4068–4076.
- Wang, H., Deng, Z., Feng, B., Ma, H., Xia, Y., 2017. An adaptive Kalman filter estimating process noise covariance. *Neurocomputing* 223, 12–17.
- Wang, S., Jiang, H., Qiao, Y., Jiang, S., Lin, H., Sun, Q., 2022. The research Progress of vision-based artificial intelligence in smart pig farming. *Sensors* 22 (17).
- Wanga, H., Peter, Ghani, Nasir, Kalegele, Khamisi, 2015. “designing a machine learning-based framework for enhancing performance of livestock mobile application system.” *American. J. Softw. Eng. Appl.* 4 (3), 56.
- Yin, H., et al., 2012. Analysis of factors on the error of the camera calibration. *Inform. Commun.* 1, 28–30.
- Yosinski, J., Clune, J., Bengio, Y., et al., 2014. How Transferable Are Features in Deep Neural Networks? In *Proceedings International Conference on Neural Information Processing Systems*. MIT Press, Kuching, Malaysia, pp. 3320–3328.