

Article

Modeling of a Hardware and Software System for Non-Invasive Monitoring of the Feeding Behavior of Farm Animals

Oleg Ivashchuk ^{1,*} , Zhanat Kenzhebayeva ^{1,*} , Alexei Zhigalov ² , Moldir Allaniyazova ¹ ,
Gulnara Kazyeva ¹ , Kaiyrbek Makulov ¹ , Vyacheslav Fedorov ³ and Olga Ivashchuk ³

- ¹ Department of Computer Science, Faculty of Science and Technology, Caspian University of Technology and Engineering (KUTI) Named After Sh. Yessenov, Microdistrict 32, Aktau 130000, Kazakhstan; moldir.allaniyazova@yu.edu.kz (M.A.); gulnara.kazyeva@yu.edu.kz (G.K.); kaiyrbek.makulov@yu.edu.kz (K.M.)
- ² Individual Entrepreneur Zhigalov A.A. (neural.dev), Agmashenebeli Str., 1, Batumi 6004, Georgia; jgaloff@gmail.com
- ³ Department of Information and Robotic Systems, Belgorod State National Research University, 85, Pobedy St., Belgorod 308015, Russia; vyachesfedorov@gmail.com (V.F.); olga.ivashuk@mail.ru (O.I.)
- * Correspondence: oleg.ivashchuk@yu.edu.kz (O.I.); zhanat.kenzhebayeva@yu.edu.kz (Z.K.); Tel.: +7-705-994-56-82 (O.I.); +7-(701)-125-61-23 (Z.K.)

Abstract

This paper presents the design of a hardware–software system for non-invasive automated monitoring of feeding behavior in livestock with biometric identification of individual animals. Neural network models for animal identification from images and individual recognition have been developed and trained. A solution is proposed to address the challenge of acquiring a sufficient number of personalized animal images for training the identification neural network. A transfer learning approach is introduced for pig identification, where the network is first trained on a large-scale dataset of more than three million human face images obtained from open sources and subsequently fine-tuned by training the upper layers on a significantly smaller dataset consisting of 5610 pig face images. Experimental results demonstrated the high effectiveness of the system: the Top-1 identification accuracy reached 95.1%, while the ROC AUC in open-set recognition tasks achieved 0.95. The processing time per frame on an NVIDIA RTX 4090 GPU was 1.4 ms (724 FPS).

Keywords: internet of things; automation of animal monitoring; smart feeder; neural network; detection; face recognition; animal recognition; pig identification



Academic Editor: Xavier Fernando

Received: 7 January 2026

Revised: 9 February 2026

Accepted: 13 February 2026

Published: 18 February 2026

Copyright: © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

1. Introduction

Global trends in livestock farming increasingly focus on non-invasive animal health monitoring, where observation includes tracking movement patterns and behavioral indicators for the early detection of diseases and abnormal conditions. Among these indicators, feeding behavior—namely the frequency and duration of food and water intake—plays a key role in assessing animal health and productivity.

Currently, video-based monitoring combined with automated image analysis is considered one of the most informative non-invasive approaches in livestock research. Existing methods include object detection using bounding boxes [1,2], segmentation of individual animals [3], trajectory tracking [4], image-based body weight estimation [5], and recognition of behavioral and emotional states such as aggression [6].

An essential component of intelligent monitoring systems is individual animal identification, which enables personalized behavioral analysis. Pigs housed in dense, dynamic groups are of particular interest for the development of robust identification algorithms due to their low inter-individual geometric variability. Early approaches proposed identification based on back hair patterns [7]; however, such methods are not applicable to animals of the same breed with similar appearance. Consequently, most recent studies employ face-based pig identification using neural networks [8,9].

A review of the scientific literature indicates that, to the best of the authors' knowledge, existing methods predominantly operate in a closed-set identification paradigm, where all individuals must be known during training. In contrast, open-set identification, which allows previously unseen animals to be detected and incorporated into the system without retraining the neural network, remains largely unexplored for pig face recognition. This limitation significantly reduces the practical applicability of current solutions in real farming environments with dynamically changing animal groups. Unlike traditional closed-set classification methods, which require complete retraining of the neural network whenever a new animal is added, the proposed open-set approach enables dynamic registration of new individuals. This is achieved by simply computing a reference feature vector (centroid) for the new animal, eliminating system downtime and making the approach scalable for real-world farming conditions with a continuously changing herd. In this study, a biometric neural network-based open-set pig face identification method is proposed.

A systematic comparison of recent livestock biometric identification approaches published between 2023 and 2026 is presented in Table 1.

Table 1. Comparison of Recent Livestock Identification Methods.

Reference	Year	Species	Identification Paradigm	Feature Representation	Supports New Individuals Without Retraining	Dataset Size	Real-Farm Validation
Hansen et al. [8]	2018	Pig	Closed-set	CNN + softmax classifier	No	~1000 images	Partial
Yin et al. [9]	2024	Pig	Closed-set	CNN + attention mechanism	No	~3000 images	Yes
Zhuang et al. [10]	2023	Pig	Closed-set	CNN-based classifier	No	~2000 images	Yes
Li et al. [11]	2022	Pig	Closed-set	YOLO-based detection + classifier	No	Not specified	Lab
Li et al. [12]	2024	Sheep	Open-set	Metric embedding (SheepFaceNet + Li-ArcFace)	Yes	~3800 images	Yes
Hou et al. [13]	2024	Multi-species	Open-world/re-ID	Metric embedding (person re-ID adapted)	Yes	Multi-species datasets	Partial
Ma et al. [14]	2025	Pig	Open-set	Three-stage metric learning system	Yes	Not specified	Yes
Xu et al. [15]	2025	Sheep	Open-set	Cross-modal metric embedding (SheepNet)	Yes	Open & closed sets	Yes
This work	2025	Pig	Open-set	IRSE-50-like metric embedding	Yes	5610 images (53 pigs)	Yes

As shown in Table 1, the majority of existing pig face identification methods are formulated as closed-set classification problems [8–11], which assume that all individuals are known during training and require retraining when new animals are introduced. In recent years, several studies have begun exploring open-set biometric identification, primarily in sheep face recognition [12,15] and general animal re-identification frameworks [13]. However, open-set identification using pig facial images remains insufficiently studied, with only limited recent attempts reported [14]. This gap highlights the need for scalable open-set biometric identification methods specifically designed for pig farming environments with dynamically changing herd composition, which is addressed in the present study.

Feeding behavior analysis enables prediction of individual weight gain and provides valuable insights into animal health. For example, reduced visits to feeders or drinkers may indicate health problems; as reported in [16], pigs infected with influenza approach waterers approximately 30% less frequently. While convolutional neural network-based approaches for automatic recognition of pig feeding behavior have been proposed [17], they typically lack individual identification. Similarly, existing smart feeder solutions with video monitoring rely on paint markings rather than biometric identification [10].

The main objective of this study is to develop a hardware–software system for non-invasive automated monitoring of livestock feeding behavior, implementing a neural network-based open-set biometric identification method using facial images captured by cameras.

2. Functional Model of the Software and Hardware System for Noninvasive Monitoring of the Feeding Process of Farm Animals

During the development of the structure and prototype of the software and hardware system for non-invasive monitoring of the feeding process of farm animals, its main functions were identified and analyzed in detail. These functions form the foundation of the system, enabling comprehensive monitoring and control of animal feeding behavior, with a focus on individual tracking and data-driven decision-making. The primary functions of the system include the following:

Detection (highlighting) of animals in images: This involves identifying the location of animals within an image or video frame. The system employs advanced neural network models for this task, ensuring accurate detection even in crowded enclosures.

Identification of animals by images (in an open set): This function enables the system to recognize and differentiate individual animals based on their unique features. Identification is particularly challenging in an open-set scenario, where new, previously unseen individuals may be encountered.

Automatic counting of the number of individuals in the enclosure: This feature provides an automated method to tally the total number of animals present in the monitored area, which is essential for managing large groups of livestock.

Calculating the number of approaches each animal makes to the feeder and waterer: By tracking individual movements, the system records how often animals visit feeding and drinking stations, offering valuable insights into their behavioral patterns.

Calculating the duration of food and water consumption for each animal: This function provides precise measurements of the time spent by each animal at the feeder and waterer, enabling a detailed analysis of feeding and drinking behavior.

Data analysis to determine integral indicators: The system aggregates and analyzes data to generate comprehensive indicators of food and water consumption, tracks the dynamics of each animal's feeding behavior, and identifies deviations from normative patterns. Based on this analysis, predictions of individual weight gain can also be made.

Generating recommendations for responsible personnel: The system produces actionable recommendations for farm staff, including optimizing feeding schedules and adjusting housing conditions, such as the temperature regime in the enclosure. These recommendations are based on the detected changes in feeding behavior, which may indicate discomfort or adverse environmental conditions.

Notifying operators of detected deviations: The system generates real-time notifications for responsible personnel regarding deviations from normal feeding behavior, facilitating early detection of diseases or other problematic conditions that may affect animal health.

Among these functions, the identification of individual animals is considered the most challenging to implement. This complexity arises from the need to accurately detect animals in images, generate unique feature representations, and match these representations to reference data in a robust and reliable manner.

The identification process is illustrated by the functional diagram shown in Figure 1. This process involves several interconnected stages:

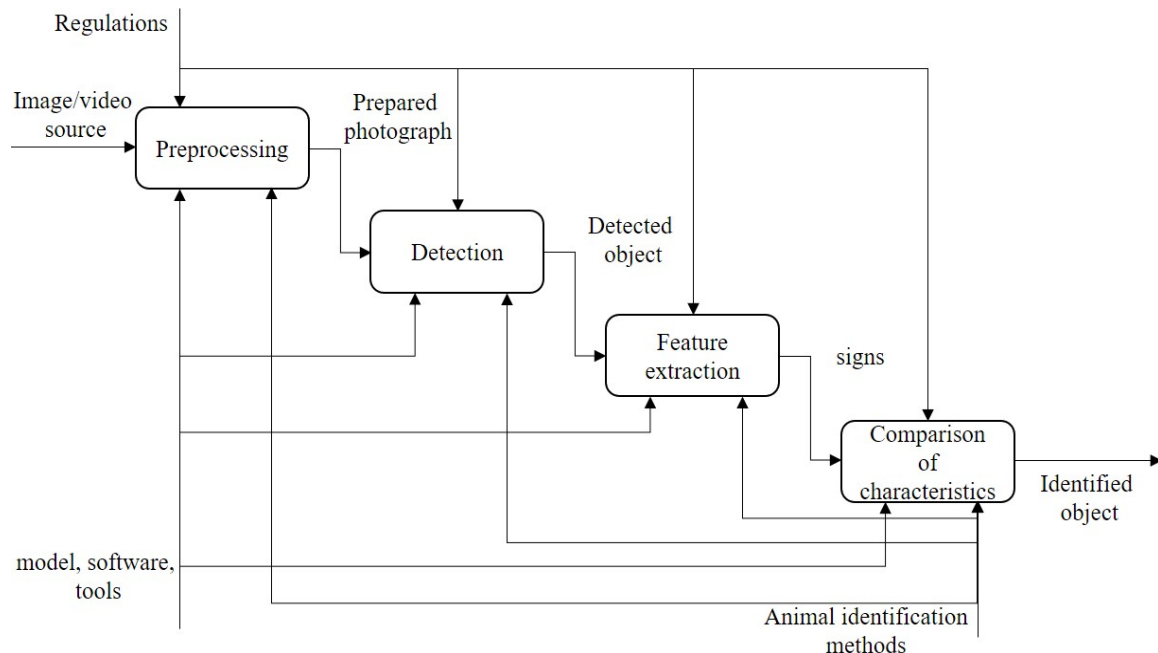


Figure 1. Functional Diagram of the Animal Identification Process.

Preprocessing: At this stage, the raw image data is converted into a format suitable for processing by the neural network responsible for detecting animals. This step ensures consistency in input data and prepares it for subsequent analysis.

Detection: Using advanced object detection techniques, the system calculates the coordinates of bounding boxes around the animals in the images. This stage isolates the relevant portions of the image, reducing noise and focusing the analysis on individual animals.

Feature vector formation: Once the animals are detected, a neural network generates a vector of distinctive features for each animal. These feature vectors encapsulate the unique visual characteristics of the detected animals, such as texture, shape, and other biometric attributes.

Feature vector comparison: The system compares the generated feature vectors with those stored in the database (DB, Database) containing reference images of known animals. The comparison involves calculating the similarity between the vectors, typically using cosine or Euclidean distance metrics. The neural network is trained to ensure that the distance between vectors of different images of the same animal is small, while the distance between vectors of different animals is large. A predefined similarity threshold determines whether the target image matches any image in the database.

The process of analyzing animal feeding behavior (including drinking behavior) is equally sophisticated and is presented in the form of a functional diagram in Figure 2. This analysis is performed based on data obtained from video cameras placed strategically within the enclosure. The key aspects of the analysis include the following:

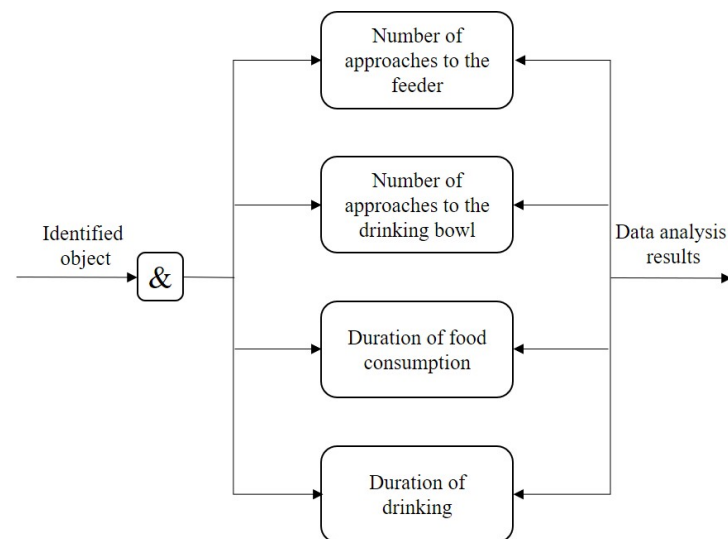


Figure 2. Functional Diagram of the Feeding Behavior Analysis Process.

Number of approaches to the feeder and waterer: For each individual, the system tracks and records the number of times the animal visits the feeding and drinking stations over a specified period. This data helps identify trends and patterns in feeding behavior.

Duration of food and water consumption: The system calculates the precise time spent by each animal consuming food and water. This metric is crucial for understanding the nutritional and hydration needs of the animals, as well as identifying any deviations from normal behavior.

Behavioral dynamics and deviations: By analyzing the data over time, the system identifies trends and detects deviations from expected behavioral patterns. For example, a sudden decrease in feeding time or frequency could indicate potential health issues, environmental discomfort, or changes in social dynamics within the group.

Predictive insights: Based on the collected data, the system generates predictive insights, such as forecasts of individual weight gain. These predictions allow farm managers to optimize feeding strategies and ensure that each animal receives adequate nutrition.

The integration of these functions creates a comprehensive system capable of monitoring and analyzing animal behavior in real time. This level of detail not only enhances the understanding of feeding patterns but also enables proactive management of livestock health and welfare. The diagrams in Figures 1 and 2 provide a clear visualization of the functional workflows, illustrating the interplay between detection, identification, data analysis, and actionable outcomes.

Feeding Behavior Event Definition and Computation of Individual Indicators

The monitoring objective of the proposed system is to transform a continuous video stream into individual-level feeding/drinking events and aggregated behavioral indicators. This is achieved by combining (1) robust face detection, (2) biometric identification in an open-set setting, and (3) event segmentation around the feeder and waterer entrances.

Regions of interest (ROIs) are defined for the feeder and waterer zones based on the fixed camera geometry and the known feeder layout. A feeding/drinking “visit” event is defined as a continuous temporal interval, during which (a) an animal face is detected within the corresponding ROI and (b) the same individual identity is assigned with stable confidence. To reduce spurious events caused by short-term occlusions, a visit is initiated only if the ROI presence is maintained for at least N consecutive frames (temporal persistence criterion), and it is terminated when the face is absent from the ROI for more than M consecutive frames (gap criterion). In practical deployment, N and M are chosen to

be small to preserve sensitivity while suppressing frame-level noise; their role is to provide simple temporal debouncing rather than to alter the underlying recognition decisions.

Identity assignment for an event is based on the metric embedding produced by the identification network. For each detected face, an embedding vector is computed and compared to the database of reference embeddings (centroids) using a similarity measure (cosine similarity or Euclidean distance). If the best match exceeds a predefined similarity threshold, the sample is assigned to the corresponding known individual; otherwise, the sample is labeled as “unknown”. In the open-set operational mode, a new individual can be registered by accumulating several embeddings captured under typical conditions and computing a reference centroid for this new identity. This procedure does not require re-training of the neural network and enables scalable operation under dynamically changing herd composition.

For each individual i and each station $s \in \{\text{feeder, waterer}\}$, the system computes the following primary indicators over a chosen monitoring interval (e.g., per hour or per day): (1) the number of visits $V(i,s)$, defined as the count of segmented events; (2) total duration $D(i,s)$, defined as the sum of event lengths; (3) mean visit duration, defined as $D(i,s)/V(i,s)$ for $V(i,s) > 0$; (4) temporal dynamics, represented as time series of $V(i,s)$ and $D(i,s)$ in sliding windows. These indicators are stored in the system database together with timestamps and confidence values, enabling retrospective analysis and automated detection of deviations from individual baselines.

The design of the feeder ensures that only one animal can access each compartment at a time, which substantially simplifies event segmentation and reduces identity ambiguity at the entry zone. Consequently, the computed behavioral indicators can be reliably attributed to a single individual and used for individualized monitoring and subsequent decision support.

3. Method for Biometric Identification of Animals in an Open Set

To accurately identify an animal by its image, it is first necessary to detect the animal within the image. In the context of scientific and technical literature on neural networks, this process is commonly referred to as “detection,” which involves the use of bounding boxes to locate objects within an image. Detection is a critical first step in the broader task of identification, as it isolates the relevant area of the image for subsequent analysis. This study proposes a method for identifying pigs by their “faces,” emphasizing the need for robust detection techniques before the identification phase. The proposed approach is integrated into the automated animal health monitoring system, where images are continuously received from cameras monitoring livestock in real time.

The specific task of detecting pig “faces” in images—particularly in challenging scenarios such as dense groups where multiple animals are in close proximity—requires advanced methods. This issue has been addressed in several studies co-authored by the authors of this article [18,19]. Their method achieves a detection accuracy of $mAP = 94.05\%$, which represents a significant improvement over the accuracy levels reported in similar studies (e.g., in [11]). This high detection accuracy is a crucial foundation for the subsequent identification stage, as errors in detection can propagate through the system, adversely affecting identification accuracy and overall system reliability. In the present study, this method was further improved, resulting in an increase in detection accuracy to $mAP@50 = 99.5\%$, where mAP denotes mean Average Precision, which provides a reliable foundation for subsequent identification.

Building on this robust detection framework, the authors developed a specialized neural network model to identify pigs by their “faces.” The model is designed to process images of pig “faces” that have been detected using the aforementioned detection

method. For effective implementation within the monitoring system, the authors propose a neural network with a structure similar to IRSE-50 (Improved Residual Squeeze-and-Excitation Network) [20]. This architecture is well-suited for image recognition tasks and is particularly advantageous when combined with a transfer learning approach.

The training process of the network is divided into two stages to optimize both performance and computational resource utilization. In the first stage, the network is pre-trained on a large-scale human face dataset comprising more than 3 million images of 85,000 individuals. This pre-training allows the network to learn general-purpose visual representations, such as edges, textures, and shapes, which are transferable across different domains. In the second stage, the network is fine-tuned on a smaller, domain-specific dataset of pig face images, consisting of approximately 5610 images from 53 individual animals. The increased dataset size compared to earlier prototypes substantially enhanced the model's generalization capability. During fine-tuning, the upper layers of the network are adapted to capture subtle, species-specific facial features of pigs, while the deeper layers responsible for extracting low-level and generic features are frozen and excluded from further training.

The goal of the identification module in this work is not generic visual categorization but biometric discrimination of individuals, i.e., learning an embedding space where samples of the same individual are close and samples of different individuals are far apart. Large-scale human face datasets are specifically curated for identity-level separation and typically contain high intra-class variability in pose, illumination, occlusion, and image quality, which matches the target challenges in pig-face identification under farm conditions. In addition, our training objective (SubCenterArcFace/AdaFace-style margin losses) is widely adopted in face-recognition pipelines and is naturally aligned with face-domain pretraining. Therefore, face pretraining provides a strong prior for metric embedding learning and stabilizes convergence when the target dataset is relatively small.

At the same time, we acknowledge that our current experiments primarily demonstrate that transfer learning outperforms training from scratch, while a systematic comparison with other common pretraining strategies (e.g., ImageNet classification pretraining or self-supervised representation learning) was not included in the present study. Such a comparison will be addressed in future work to further quantify the relative benefits of different pretraining sources for livestock biometric identification.

Transfer learning is a widely used technique in neural network development because it allows models to leverage the knowledge gained from a large, readily available dataset to perform well on a smaller, more specialized target dataset. This approach is particularly beneficial in cases like this, where obtaining a large dataset of personalized pig “face” images is technically challenging. The need for manual labeling of training data, coupled with the inherent difficulties of distinguishing pigs due to their similar features, dirt accumulation, and frequent clustering in dense groups, makes transfer learning a highly justified choice. By freezing the deep layers, which extract general image features, and focusing training on the top layers to refine species-specific characteristics, the authors ensure that the network achieves high accuracy without overfitting the smaller dataset.

As the loss function, the authors developed and applied a modified SubCenterAdaFace loss. This loss function combines the advantages of two state-of-the-art approaches: Sub-Center ArcFace and AdaFace.

Sub-Center (sub-centroids): Instead of using a single centroid for each class (pig), K sub-centroids ($K = 3$) are employed. This allows the model to effectively handle intra-class variability, where the same animal may appear differently depending on pose (e.g., frontal or profile views) or illumination conditions. During training, only the sub-

centroid closest to the current sample is selected, which improves convergence on noisy data [21].

AdaFace (adaptive margin): This component dynamically adjusts the angular margin based on image quality, estimated via the feature vector norm. For high-quality images, the margin is increased to encourage the learning of more discriminative features, whereas for blurred or low-quality frames, the margin is reduced to prevent overfitting to unreliable data [22].

The combination of these two mechanisms significantly improves recognition robustness under challenging real-farm conditions, including contamination, animal motion, and suboptimal lighting.

Fine-tuning the network on the pig “face” dataset is expected to yield optimal identification accuracy, but this process requires careful experimental adjustments. Key parameters that need to be optimized include the number of “frozen” layers, the method for minimizing the loss function (typically a variant of gradient descent), the learning rate for gradient descent, and other hyperparameters governing the training process. These adjustments ensure that the network achieves the best balance between generalization and specificity, allowing it to perform well in real-world scenarios.

The creation of the pig face dataset involved significant technical and logistical challenges, as it required collecting a sufficiently large, diverse, and representative set of images under real farming conditions. The total dataset comprises 5610 images. Of these, 2000 images of 23 pigs were obtained from publicly available online sources, while the remaining 3610 images of 30 individual animals were collected by the authors through direct photo and video recording at the Agricultural Production Cooperative (APC) “Kolkhoz named after Gorin” located in the Belgorod region.

To ensure training stability and class representativeness, a filtered subset of 4907 images corresponding to 53 individual pigs was selected from the full dataset, excluding classes containing fewer than 15 images. The distribution of the number of samples per individual is illustrated in Figure 3. The resulting class distribution is inherently imbalanced, ranging from 15 to 357 images per individual, which reflects realistic farm operating conditions.

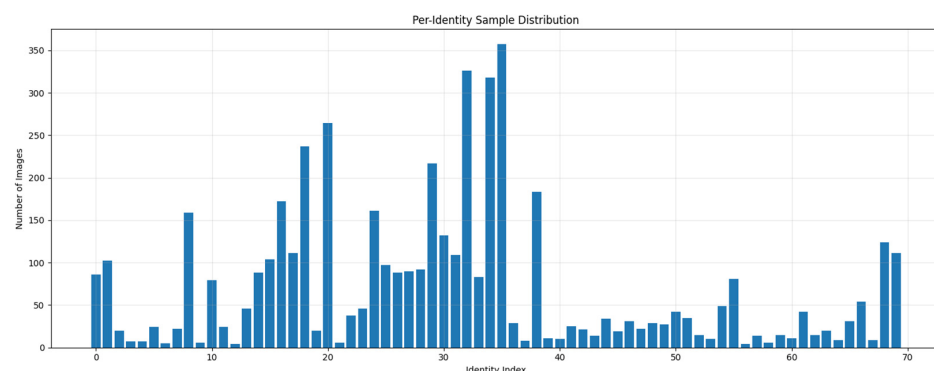


Figure 3. Distribution of the number of images per individual pig in the selected dataset.

Model performance was evaluated using an independent test set comprising at least 12% of the data. To ensure strict reproducibility, the split between training and test sets was performed using a fixed random seed. This approach, combined with the size and diversity of the independent test set, ensures the statistical reliability of the reported evaluation metrics and is consistent with standard practices for training deep neural network architectures.

All images and videos used for training and evaluation were acquired under real commercial farming conditions without any staged setups or artificial constraints, ensuring ecological validity of the collected data.

Figure 4 presents a general view of the pig pen where the data acquisition was conducted, illustrating the typical environment in which the animals were observed.



Figure 4. General View of the Pen Where Pigs Were Filmed Using Photo and Video Cameras.

Figure 4 illustrates the real operating environment in which the proposed system is deployed. This environment defines key visual challenges for identification and behavior analysis, including variable illumination, partial occlusions, animal clustering, and contamination of visual surfaces.

The recording process involved capturing images and videos under various conditions to ensure that the dataset captured a wide range of pig appearances and behaviors. Table 2 lists the specific methods used for recording, including details on camera positioning, lighting conditions, and strategies for isolating individual pigs within groups. Table 2 summarizes the practical data acquisition strategies applied during photo and video recording, highlighting their advantages and limitations under real commercial farming conditions. The authors prioritized creating a dataset that would reflect the challenges encountered in real-world applications, such as variations in lighting, partial occlusions, and the presence of dirt or debris on the pigs.

Table 2. Data acquisition strategies and recording conditions used for constructing the pig face image dataset under real-farm conditions.

Way	Peculiarities
Taking photos manually	It is advisable to mark animals with paint whose photographs have already been taken to avoid repeated imaging.
Video recording using cameras attached to ceiling beams. Next, it is necessary to track the movement trajectory of each pig and select frames containing the animal’s “face”.	Efforts should be made to avoid identification errors when reviewing video footage. Errors may arise due to similarities between pigs, dirt on their bodies, and their rapid movement.
Filming pigs during herding using three video cameras mounted above and on the sides of the passage.	It is advisable to illuminate the passage. Since the pigs’ “faces” are mostly not visible when they move along the passage, this method yields only a small number of images.

4. Description of the Prototype Software and Hardware System for Non-Invasive Monitoring of the Pig Feeding Process

The developed prototype of the hardware and software system represents a comprehensive solution for automated monitoring of the feeding process of agricultural animals. This system addresses one of the key challenges in modern livestock management—ensuring accurate monitoring and control of animal feeding behavior to improve productivity and animal welfare. The central component of the system is an advanced feeder, specifically designed to

combine functionality and durability in demanding agricultural environments. The feeder is equipped with two dedicated compartments for feed and water, ensuring separation of functions for optimal monitoring.

Each compartment is engineered to accommodate only one animal at a time, a critical feature that ensures the precise operation of the identification system. By preventing overlapping and merging of “faces” in video recordings, the system can reliably identify individual animals and track their feeding behavior with a high degree of accuracy.

The video surveillance system is implemented using high-quality IP cameras (Hikvision DS-2CD2T43G2-2I/4I) with a focal length of 2.8 mm and a maximum resolution of 2688×1520 . The cameras are installed at a height of 1.5 m and at a distance of 1.5 m from the feeding area, with an installation angle of 45° , providing sufficient image clarity for reliable animal identification.

The cameras are housed in durable protective enclosures, ensuring stable operation under harsh livestock farming conditions characterized by dust, moisture, and temperature fluctuations. To enhance functionality, the cameras are equipped with infrared illumination, enabling reliable monitoring under low-light conditions, such as at night or in poorly lit environments. This feature ensures continuous 24/7 operation, allowing uninterrupted data acquisition.

For continuous video streaming from a 4-megapixel camera, the required communication bandwidth is approximately 5–10 Mbps per camera, depending on the compression settings and frame rate.

Figure 5 illustrates the schematic design of the hardware component of the feeder, highlighting the integration of non-invasive monitoring technologies specifically tailored for tracking the feeding behavior of pigs. The system’s hardware is complemented by advanced processing capabilities. Video streams captured by the cameras are processed on a dedicated server equipped with a high-performance processor and a powerful graphics accelerator, which are essential for running sophisticated neural network algorithms in real time. The server infrastructure also includes large-capacity data storage devices to archive observation results over extended periods. This long-term data storage capability facilitates trend analysis and supports evidence-based decision-making in livestock management.

The software architecture of the system is modular, offering flexibility and ease of updates. Key components include a database that stores reference images of pigs’ “faces,” biometric parameter vectors, and comprehensive observation histories. This database serves as the foundation for the system’s neural network module, which implements algorithms for detecting and identifying animals. Operating in real time, the module ensures continuous analysis of the video feed, with a high level of precision. Additionally, an analytical subsystem processes the gathered data, producing detailed reports on animal behavior, feeding patterns, and any detected anomalies. These reports provide valuable insights for livestock managers, enabling them to take timely corrective actions to optimize feeding strategies and ensure animal health.

Special attention was devoted to the development of user interfaces to make the system accessible and functional for various categories of personnel. Automated workstations (AWS) are provided for system administrators, livestock operators, and other staff, each configured with software tailored to their specific roles and levels of authority. For example, system administrators can access advanced configuration tools, while livestock operators are provided with simplified interfaces for day-to-day monitoring and management tasks. The intuitive design of these interfaces minimizes the training required for staff and ensures efficient system operation.

A mobile application was also developed as part of the system, aimed at enhancing operational flexibility and on-site functionality. The application, designed for the Android

platform, allows livestock facility staff to identify pigs directly in production environments using the camera of a mobile device. The identification process is seamless and occurs in real time: After capturing an animal's "face" using the smartphone's camera, the image is processed by a built-in neural network optimized for mobile devices. The neural network performs rapid analysis, displaying the animal's unique identification number and current status on the device screen. The application also allows users to add new animals to the database by capturing a series of reference photographs on-site. Integration between the mobile application and the central system is achieved via a secure connection, ensuring real-time data synchronization and up-to-date information across all system components.

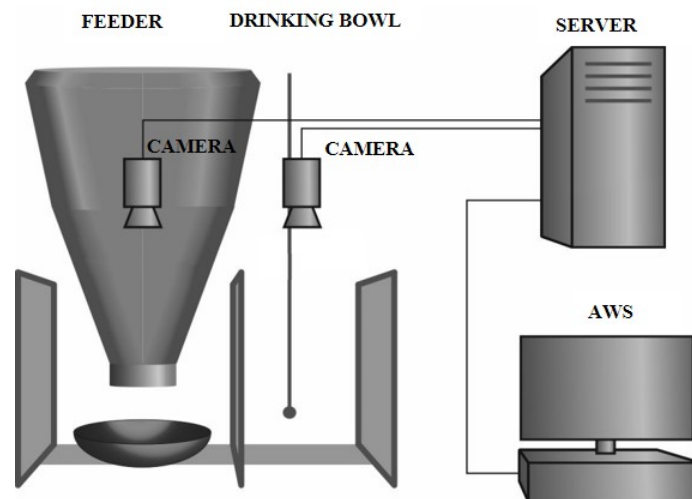


Figure 5. Schematic of the Hardware Part of the Feeder Equipped with Non-Invasive Monitoring Tools for the Pig Feeding Process.

The system implements a wide range of functions for monitoring the feeding process, including continuous tracking of animal visits to the feeder, measurement of feeding duration, analysis of feed and water consumption rates, and detection of irregularities in feeding behavior. A notable feature of the system is its ability to automatically generate notifications for personnel when non-standard situations are detected, such as unusual feeding patterns or signs of potential health issues. These notifications ensure timely intervention and contribute to maintaining the overall health and productivity of the livestock.

The developed system is highly scalable and can be adapted for use in livestock facilities of various sizes, from small farms to large industrial complexes. Its modular architecture not only facilitates easy upgrades and component replacement but also enables integration with existing livestock management systems, offering a seamless transition for facilities looking to modernize their operations.

To ensure reliability under challenging conditions, all components of the system are designed to withstand the demanding environment of livestock facilities. The hardware is vandal-proof, resistant to aggressive substances, and compliant with IP66 protection standards, ensuring durability and consistent performance. Additionally, the system is equipped with uninterruptible power supplies (UPS), guaranteeing operation during brief power outages. This feature ensures data integrity and uninterrupted monitoring, even in areas where power supply is unstable.

Figure 5 demonstrates how the physical design of the feeder directly supports the computational pipeline. By restricting access to a single animal per compartment, the hardware layout simplifies identity assignment and enables unambiguous attribution of feeding events to individual animals.

5. Experimental Results

To validate the effectiveness of the proposed system, quantitative evaluation of the detection and identification modules was conducted using independent test datasets that were not used during training.

5.1. Detection Performance Evaluation

To assess the robustness of the detection algorithm, an independent test dataset consisting of 768 images, which was not used during training, was employed. The results are summarized in Table 3.

Table 3. Quantitative performance metrics of the detection module on the independent test set.

Metric	Value	Description
Precision	1.000	No false positives were produced by the model
Recall	0.997	Only a negligible number of objects were missed
mAP@50	0.995	High average precision at the standard IoU threshold
mAP@50–95	0.970	Highly accurate bounding box localization

Achieving Precision = 1.000 with Recall = 0.997 indicates highly reliable separation of pig faces from the background, providing a robust basis for subsequent identification.

5.2. Identification Performance Evaluation

The performance of the identification module was evaluated using an independent test dataset. The key quantitative metrics are presented in Table 4.

Table 4. Quantitative performance metrics of the identification module on the independent test set.

Metric	Value	Interpretation
Top-1 Accuracy	95.09%	Probability of correct individual identification on the first attempt
Top-5 Accuracy	100.00%	Correct individual consistently appears among the top five candidates
F1 Score	0.842	Balanced measure of precision and recall
Inference Speed	724 FPS	Processing speed (frames per second) on an NVIDIA RTX 4090 GPU

The Top-1 accuracy reached 95.09%, while Top-5 accuracy achieved 100%, indicating that the correct identity is consistently retained among the most probable candidates. This property enables effective application of temporal smoothing and decision aggregation across video sequences, resulting in near-perfect system reliability in practical deployment scenarios.

It should be noted that the metrics reported in Table 3 correspond to the closed-set identification scenario, where all test identities are known to the system and a fixed similarity threshold is used. In this setting, Top-1 and Top-5 accuracy metrics are well-defined and commonly used to evaluate identification reliability, while the reported F1-score reflects the system performance under a fixed operating point.

The reported inference speed of 724 FPS corresponds to single-frame inference (batch size = 1) on an NVIDIA RTX 4090 GPU and includes both face detection and feature embedding extraction stages. This value represents an upper-bound performance estimate under high-performance hardware conditions and is provided to demonstrate the real-time capability of the proposed system. For improved interpretability, this throughput corresponds to an average processing latency of approximately 1.4 ms per frame. Such latency enables real-time operation even when multiple video streams or feeders are processed

concurrently on a single GPU. Direct comparison of FPS values with existing studies is often unreliable due to differences in hardware platforms, input image resolution, network architectures, and evaluation protocols. Therefore, inference latency per frame is reported as a more hardware-agnostic indicator of computational efficiency and practical deployability.

5.3. Demonstration of System Outputs for Feeding and Drinking Monitoring

The behavioral indicators computed in this section directly correspond to the functional workflow illustrated in Figure 2, transforming conceptual monitoring stages into measurable system outputs.

Beyond detection and biometric identification accuracy, the proposed system produces individual-level monitoring outputs by assigning identities to feeder and waterer visit events and aggregating behavioral indicators over time. Each recorded event is stored in the database as a structured entry containing the following: (1) timestamp of event start and end, (2) station type (feeder or waterer), (3) assigned identity (known ID or “unknown”), (4) similarity/confidence values associated with the decision, and (v) auxiliary metadata (camera ID, frame indices).

From these event logs, the system computes individual behavioral summaries such as the number of visits and total consumption duration per individual and per station within a specified interval. These outputs directly correspond to the monitoring objectives illustrated in Figure 2 and represent the final operational product of the hardware–software system. The feeder geometry (single-animal access per compartment) ensures that each event corresponds to one individual, enabling unambiguous attribution of feeding/drinking behavior.

The practical relevance of the open-set identification paradigm is particularly evident at the system-output level: when an unregistered animal appears in the monitored group, its events are recorded as “unknown” rather than being forcibly assigned to an existing identity. After registration, subsequent events are consistently attributed to the new identity via its reference centroid, without retraining the network. Therefore, the monitoring pipeline remains operational under changing herd composition and produces complete longitudinal behavior records for all registered animals.

5.4. Open-Set Identification Analysis

To evaluate the system’s ability to reject unknown individuals, an open-set recognition test was conducted. The ROC curve is shown in Figure 6.

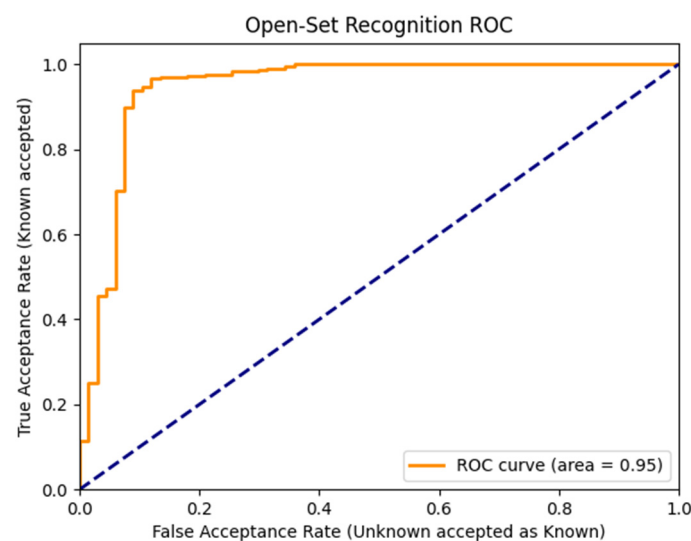


Figure 6. ROC curve for open-set recognition.

The system achieved an AUC of 0.95, indicating a high discriminative capability between known and unknown animals. At a false acceptance rate (FAR) of 1%, the true acceptance rate (TAR) exceeds 80%.

5.5. Threshold Optimization and F1-Score

To achieve a balance between identification precision and recall in the open-set recognition scenario, the effect of the similarity threshold on the final performance metrics was analyzed. Unlike the closed-set evaluation presented in Table 3, this analysis considers the open-set setting, where both known and unknown individuals may appear, and the decision threshold plays a critical role in controlling false acceptances and false rejections.

Figure 7 illustrates the dependence of the F1-score on the selected similarity threshold. The maximum value of $F1 = 0.976$ is achieved at a threshold of 0.30. This value corresponds to the optimal operating point identified during off-line threshold optimization and was selected for real-world system deployment.

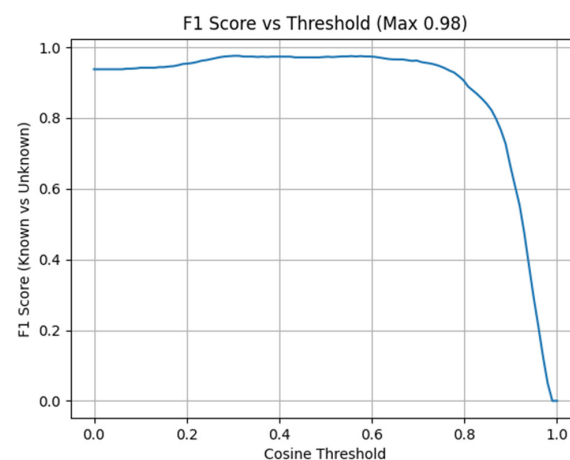


Figure 7. F1-score as a function of the similarity threshold.

The difference between the F1-score reported in Table 3 and the maximum F1-score observed in Figure 7 is explained by the different evaluation protocols: fixed-threshold closed-set testing versus threshold-optimized open-set analysis.

5.6. Convergence Analysis Under Transfer Learning

The effectiveness of knowledge transfer from human face representations is confirmed by a comparison of training convergence curves (Figure 8).

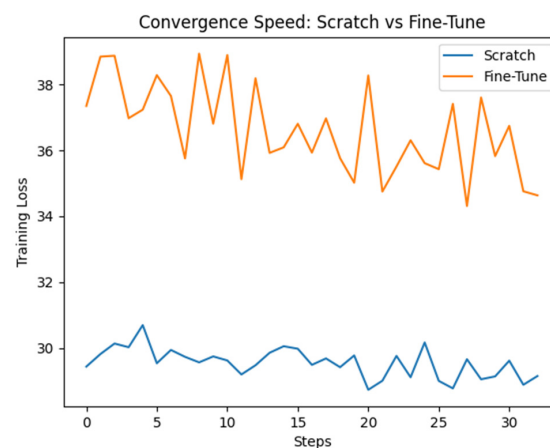


Figure 8. Convergence curves for training from scratch and fine-tuning.

Training from scratch exhibits loss stagnation during the first five epochs, whereas fine-tuning demonstrates immediate performance improvement, which is critical when operating with limited domain-specific data.

5.7. Error Analysis

Visualization of the confusion matrix (Figure 9) and analysis of challenging cases indicate that the primary sources of errors are severe motion blur and extreme occlusions, such as contamination of the camera lens.

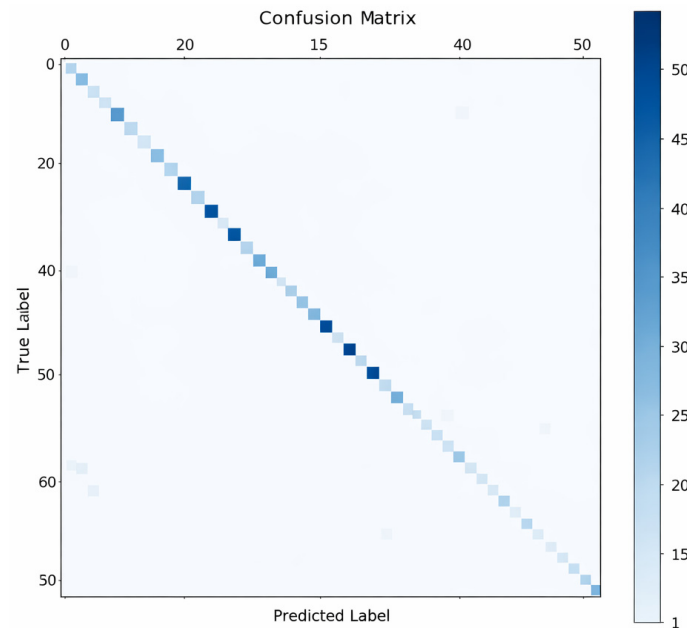


Figure 9. Confusion matrix for the identification process.

A manual inspection of the independent test set indicates that frames affected by severe motion blur constitute approximately 6–8% of the total samples. These cases are primarily associated with rapid head movements near the feeder entrance and represent extreme conditions rather than typical operating scenarios.

Representative examples of such challenging recognition scenarios are shown in Figure 10.

The examples shown in Figure 10 correspond to extreme operating conditions rather than typical system behavior and are provided to illustrate the main sources of residual errors under real-farm conditions.

While Figure 10 presents representative examples of recognition failures, a comprehensive assessment of the proposed system also requires qualitative analysis of successful identification cases. Figure 11 illustrates examples of correct recognition selected from the independent validation dataset.

Despite challenging conditions such as low illumination, partial facial occlusion, and extreme head poses, the model consistently assigns the correct identity with high confidence. From an analysis perspective, these examples indicate that the learned embedding captures stable identity-related facial structures rather than transient visual artifacts caused by pose variation, lighting changes, or partial occlusions. This qualitative evidence complements the quantitative evaluation and confirms the robustness of the proposed open-set biometric identification approach under realistic farm operating conditions.

Taken together, Figures 10 and 11 provide a balanced qualitative assessment of the proposed system, illustrating both rare failure modes and typical successful operation under real-farm conditions.



Figure 10. Examples of failure cases and challenging recognition scenarios.



Figure 11. Examples of successful pig face recognition under challenging conditions. The figure presents eight representative cases in which the proposed model correctly identifies individual pigs despite adverse conditions, including low illumination, partial facial occlusion, and extreme head poses. For each image, the ground-truth identity (True), predicted identity (Pred), and confidence score (Conf) are shown. All examples are taken from the independent validation dataset and were not used during training.

6. Conclusions

This study presented a hardware–software system for non-invasive monitoring of feeding and drinking behavior in livestock, integrating computer vision modules with a dedicated feeder design that supports reliable individual-level event attribution. The proposed approach combines face detection, biometric identification in a metric-embedding space, and an open-set operating mode that enables registration of previously unseen animals without retraining the neural network.

Quantitative experiments demonstrated strong performance of the perception pipeline on independent test data. The detection module achieved high accuracy (Precision = 1.000, Recall = 0.997, mAP@50 = 0.995, and mAP@50–95 = 0.970), providing a robust basis for downstream identification. In closed-set identification testing, the system reached Top-1 accuracy of 95.09% and Top-5 accuracy of 100%, indicating stable identity ranking for known individuals. In open-set recognition analysis, the ROC AUC reached 0.95, confirming a high capability to separate known and unknown animals. Threshold optimization further identified an operational point (similarity threshold = 0.30) that maximizes the F1-score in the open-set setting (F1 = 0.976), supporting practical deployment where false acceptances and false rejections must be balanced.

The system is designed to deliver monitoring outputs that directly match the targeted behavioral indicators. By assigning identities to segmented feeder and waterer visit events, the software produces individual-level metrics such as the number of approaches and total consumption duration over time, enabling personalized monitoring and deviation detection. The open-set paradigm ensures that the monitoring pipeline remains functional under dynamic herd composition: unknown individuals are explicitly rejected and can be registered via centroid formation without model downtime.

Real-time feasibility was demonstrated by an inference throughput of 724 FPS (approximately 1.4 ms per frame) on an NVIDIA RTX 4090 GPU for single-frame processing, including both detection and embedding extraction stages. While FPS values are hardware-dependent, the reported per-frame latency indicates that the proposed pipeline can support real-time multi-stream operation in farm environments.

The main limitations of the present study include the absence of a systematic comparison of alternative pretraining sources (e.g., ImageNet or self-supervised strategies) for metric embedding learning, as well as the dependence of performance on image quality under extreme conditions such as severe motion blur and strong occlusions. Future works will focus on expanding such comparisons, improving robustness under challenging visual conditions, and validating long-term stability of behavioral monitoring outputs in extended farm deployments.

Overall, the proposed system provides a practical and scalable solution for precision livestock farming by combining open-set biometric identification with continuous video-based monitoring of feeding behavior, contributing to improved animal management, welfare, and productivity.

Author Contributions: Conceptualization, O.I. (Oleg Ivashchuk) and V.F.; methodology, O.I. (Olga Ivashchuk) and V.F.; software, A.Z. and K.M.; validation, O.I. (Oleg Ivashchuk), A.Z. and K.M.; formal analysis, O.I. (Oleg Ivashchuk); investigation, Z.K., M.A. and G.K.; resources, M.A. and G.K.; data curation, Z.K.; writing—original draft preparation, O.I. (Oleg Ivashchuk); writing—review and editing, V.F. and O.I. (Olga Ivashchuk); visualization, M.A.; supervision, O.I. (Oleg Ivashchuk); project administration, O.I. (Oleg Ivashchuk); funding acquisition, V.F. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Science Committee of the Ministry of Science and Higher Education of the Republic of Kazakhstan (Grant No. AP23489215). The APC was funded by the same grant.

Informed Consent Statement: This study involved non-invasive video monitoring of farm animals conducted under standard commercial farming conditions. No physical contact, manipulation, invasive procedures, or behavioral interventions were applied to the animals. The data acquisition process did not cause stress, pain, or discomfort to the animals and did not interfere with their normal feeding behavior. All observations were performed in compliance with applicable national and institutional guidelines for the ethical treatment of animals in agricultural research. According to local regulations, formal ethical approval was not required for purely observational, non-invasive monitoring studies conducted within routine livestock management practices.

Data Availability Statement: The dataset of pig face images used for training and evaluation in this study is publicly available at the Hugging Face repository: https://huggingface.co/datasets/AlexeyZhigalov/Pig_classes (accessed on 6 January 2026). The repository includes identity labels and metadata required to reproduce all experiments reported in this paper.

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

Abbreviations

The following abbreviations are used in this manuscript:

mAP	mean Average Precision
DB	Database
IRSE-50	Improved Residual Squeeze-and-Excitation Network
APC	Agricultural Production Cooperative
UPS	Uninterruptible Power Supply
AWS	Automated Workstation
HD/Full HD	High Definition/Full High Definition

References

1. Beloglazov, P.G. Unified animal identification system: Problems, solutions, and prospects. *Eff. Livest.* **2019**, *7*, 52–53.
2. Riekert, M.; Klein, A.; Adrion, F.; Hoffmann, C.; Gallmann, E. Automatically detecting pig position and posture by 2D camera imaging and deep learning. *Comput. Electron. Agric.* **2020**, *174*, 105391. [\[CrossRef\]](#)
3. Lee, S.K. *Pig Pose Estimation Based on Extracted Data of Mask R-CNN with VGG Neural Network for Classifications*; South Dakota State University: Brookings, South Dakota, 2020; p. 4098.
4. Zhang, L.; Gray, H.; Xujiong, C.; Allinson, N. Individual Pig Detection and Tracking in Pig Farms. *Sensors* **2019**, *19*, 1188. [\[CrossRef\]](#) [\[PubMed\]](#)
5. Kashiha, M.A.; Bahr, C.; Ott, S.; Moons, C.P.H.; Niewold, T.A.; Odberg, F.O.; Berckmans, D. Automatic weight estimation of individual pigs using image analysis. *Comput. Electron. Agric.* **2014**, *107*, 38–44. [\[CrossRef\]](#)
6. Camerlink, I.; Coulange, E.; Farish, M.; Baxter, E.M.; Turner, S.P. Facial expression as a potential measure of both intent and emotion. *Appl. Anim. Behav. Sci.* **2018**, *205*, 165–174. [\[CrossRef\]](#) [\[PubMed\]](#)
7. Huang, W.; Zhu, W.; Ma, C.; Guo, Y. Weber Texture Local Descriptor for Identification of Group-Housed Pigs. *Sensors* **2020**, *20*, 4649. [\[CrossRef\]](#) [\[PubMed\]](#)
8. Hansen, M.F.; Smith, M.L.; Smith, L.N.; Salter, M.G.; Baxter, E.M.; Farish, M.; Grieve, B. Towards on-farm pig face recognition using convolutional neural networks. *Comput. Ind.* **2018**, *98*, 145–152. [\[CrossRef\]](#)
9. Yin, Z.; Peng, M.; Guo, Z.; Zhao, Y.; Li, Y.; Zhang, W.; Li, F.; Guo, X. Lightweight Pig Face Feature Learning Evaluation and Application Based on Attention Mechanism and Two-Stage Transfer Learning. *Agriculture* **2024**, *14*, 156. [\[CrossRef\]](#)
10. Zhuang, Y.; Zhou, K.; Zhou, Z.; Ji, H.; Teng, G. Systems to Monitor the Individual Feeding and Drinking Behaviors of Growing Pigs Based on Machine Vision. *Agriculture* **2023**, *13*, 103. [\[CrossRef\]](#)
11. Li, G.; Jiao, J.; Shi, G.; Ma, H.; Gu, L.; Tao, L. Fast Recognition of Pig Faces Based on Improved Yolov3. *J. Phys. Conf. Ser.* **2022**, *2171*, 012006. [\[CrossRef\]](#)
12. Li, J.; Zhang, X.; Wang, Q. Open-Set Sheep Face Recognition Based on SheepFaceNet with Li-ArcFace. *Agriculture* **2024**, *14*, 1112. [\[CrossRef\]](#)

13. Hou, S.; Huang, P.; Wang, Z.; Liu, Y.; Li, Z.; Zhang, M.; Huang, Y. OpenAnimals: Revisiting Person Re-Identification for Animals Towards Better Generalization. *arXiv* **2024**, arXiv:2410.00204v1.
14. Ma, R.; Ali, H.; Waqar, M.M.; Kim, S.C.; Kim, H. Pig Face Open Set Recognition and Registration Using a Three-Stage System. *Animals* **2025**, *15*, 691. [[CrossRef](#)] [[PubMed](#)] [[PubMed Central](#)]
15. Xu, X.; Song, H.; Pan, H.; Chen, D.; Yang, S. Round-the-Clock Accurate Sheep Face Recognition via Cross-Modal Embedding (SheepNet). *Comput. Electr. Agric.* **2025**, *239*, 110918. [[CrossRef](#)]
16. Kashiha, M.A.; Bahr, C.; Ott, S.; Moons, C.P.H.; Niewold, T.A.; Odberg, F.O.; Berckmans, D. Automatic identification of marked pigs in a pen using image pattern recognition. *Comput. Electron. Agric.* **2013**, *93*, 111–120. [[CrossRef](#)]
17. Alameer, A.; Kyriazakis, I.; Dalton, H.A.; Miller, A.L.; Bacardit, J. Automatic recognition of feeding and foraging behaviour in pigs using deep learning. *Biosyst. Eng.* **2020**, *197*, 91–104. [[CrossRef](#)]
18. Zhigalov, A.A.; Ivashchuk, O.A.; Biryukova, T.K.; Fedorov, V.I. Neural network detection methods of agricultural animals in dense dynamic groups on images. *Artif. Intell. Decis.* **2023**, *1*, 55–66.
19. Ivashchuk, O.; Atymtayeva, L.; Zhigalov, A.; Yagalieva, B.; Ivashchuk, O.; Fedorov, V. Neural network methods for the detection of farm animals in dense dynamic groups on images. *Appl. Math. Inf. Sci.* **2024**, *18*, 241–249. [[CrossRef](#)]
20. Fu, K.; Li, J.; Ma, L.; Mu, K.; Tian, Y. Intrinsic Relationship Reasoning for Small Object Detection. *arXiv* **2020**, arXiv:2009.00833. [[CrossRef](#)]
21. Deng, J.; Guo, J.; Liu, T.; Gong, M.; Zafeiriou, S. Sub-center ArcFace: Boosting face recognition by large-scale noisy web faces. In Proceedings of the European Conference on Computer Vision (ECCV), Glasgow, UK, 23–28 August 2020; pp. 741–757.
22. Kim, M.; Jain, A.K.; Liu, X. AdaFace: Quality Adaptive Margin for Face Recognition. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), New Orleans, LA, USA, 18–24 June 2022; pp. 18750–18759.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.