Computers and Electronics in Agriculture 237 (2025) 110674

Contents lists available at ScienceDirect

Computers and Electronics in Agriculture

journal homepage: www.elsevier.com/locate/compag

Review article

Advancing animal farming with deep learning: A systematic review

Zahid Ur Rahman, Mohd Shahrimie Mohd Asaari *, Haidi Ibrahim

School of Electrical and Electronic Engineering, Engineering Campus, Universiti Sains Malaysia, Nibong Tebal, 14300, Pulau Pinang, Malaysia

ARTICLE INFO

Keywords:
Deep learning
Computer vision
Animal farming
Livestock monitoring
Convolutional neural networks
You Only Look Once
Memory-based neural networks
Generative adversarial networks

ABSTRACT

Deep learning has revolutionized animal farming by enabling automated health monitoring, behavior analysis, and livestock management. This review examines the application of key deep learning architectures, including convolutional neural networks (CNNs), You Only Look Once (YOLO), memory-based neural networks (MBNNs), and generative adversarial networks (GANs), in various aspects of animal farming. These models have demonstrated success in tasks such as real-time livestock detection, disease prediction, activity monitoring, and animal identification. However, challenges such as occlusion, data scarcity, small training datasets, environmental variability, and imbalanced data remain significant barriers to model reliability and scalability. By analyzing one hundred seventeen articles following PRISMA guidelines, this review highlights recent advancements, identifies research gaps, and discusses solutions such as image augmentation, synthetic data generation, and domain adaptation. The findings emphasize the potential of deep learning to enhance precision farming while addressing critical challenges. Finally, future research directions are proposed to improve model generalization, integration with IoT-based monitoring systems, and real-time decision-making for sustainable and intelligent livestock management.

Contents

* Corresponding author.
E-mail address: mohdshahrimie@usm.my (M.S.M. Asaari).

https://doi.org/10.1016/j.compag.2025.110674
Received 21 April 2025; Received in revised form 3 June 2025; Accepted 15 June 2025
Available online 7 July 2025
0168-1699/

1. Introduction

Animal farming, also known as livestock farming, involves breeding and raising animals for food production, including meat, milk, and eggs. It plays a crucial role in global food security, providing essential nutrients and supporting the livelihoods of millions worldwide. Livestock farming encompasses various species, such as cattle, poultry, pigs, and sheep, each contributing uniquely to the agricultural economy. Farming methods range from traditional practices emphasizing sustainability and animal welfare to intensive systems focused on maximizing production efficiency (Nyamuryekungwe et al., 2023).

Over the past decade, animal farming has experienced a significant transformation from traditional manual practices to intelligent, data-driven systems powered by deep learning. Before 2010, most livestock management relied heavily on human observation, simple tools, and routine checks, which were often inefficient and error-prone. Gradual digitization began with the adoption of RFID tagging, environmental sensors, and basic machine learning techniques for tasks like weight monitoring and facial recognition. Between 2010 and 2020, with the advent of more affordable computing and improved connectivity, traditional computer vision and early machine learning models were used for basic identification and disease detection, although these systems still depended heavily on handcrafted features and lacked robustness.

From 2020 onward, the integration of deep learning enabled more advanced applications such as automated behavior analysis, disease detection, and livestock tracking. This period also saw the emergence of multimodal approaches, alongside the growth of edge computing for real-time inference in low-resource farm settings. These developments addressed critical gaps in scalability and precision, the trend toward unified platforms and explainable AI systems. The shift toward sustainable and ethical AI applications in farming has further reinforced the need for deep learning driven innovation, positioning it as a central pillar of smart agriculture.

Historically, animal farming has evolved from traditional practices to more intensive and technologically advanced Artificial Intelligence (AI) systems. The domestication of animals led to the development of various breeds optimized for specific purposes such as meat, milk, and wool production. Over time, advancements in breeding techniques and husbandry practices have improved productivity and efficiency in livestock management (Chen et al., 2023c). Traditional animal farming relied on manual inspection, basic statistical tools, RFID tags, and infrared thermography, which were labor intensive and limited in scope. These methods often missed early signs of illness or stress due to infrequent checks, lacked real-time monitoring, and provided only periodic productivity data. Manual inspections were subjective and time-consuming, while RFID tags offered basic identification without detailed health or behavior insights. Infrared thermography, though useful for detecting fever, was costly and impractical for continuous use. Modern technologies like IoT sensors, and wearable devices have revolutionized farming by enabling real-time, continuous monitoring, early detection of issues, and data driven optimization of productivity and resource use (Berckmans, 2017).

List of abbreviations	
ADA	Adversarial domain adaptation
AI	Artificial intelligence
AIS	Amodal Instance Segmentation
AP	Average precision
ASC	Animal Sound Classification
AUC	Area Under the Curve
BN	Batch normalization
CCMNN	Cyclic Consistent Migration Neural Network
CDAE	Convolutional Denoising Autoencoders
CNN	Convolutional Neural Networks
COD	Camouflaged object detection
CV	Computer vision
CycleGAN	Cycle-Consistent GAN
DL	Deep learning
GAN	Generative Adversarial Network
GAN-MAE	GAN-Masked Autoencoders
GNN	Graph Neural Networks
GPUs	Graphical processor unit
IMUs	Inertial Measurement Units
IoU	Intersection over union
LS-GAN	least Squares GAN
LSTM	Long Short-Term Memory
MAE	Mean Absolute Error
mAP	mean Average Precision
MBNN	Memory-based neural network
MDAM	Multi-Scale Domain Adaptation Module
mIoU	Mean Intersection of Union
NCTH	Non-Compressible Torso Hemorrhage
OST	Ocular Surface Temperature
PAN	Path Aggregation Network
PSNR	Peak Signal-to-Noise Ratio
PUFA	Poly-Unsaturated Fatty Acids
ReLU	Rectified linear unit
RGB	Red-Green-Blue
SIFT	Scale-Invariant Feature Transform
SLR	Systematic literature review
SPP	Spatial Pyramid Pooling
SSIM	Structural Similarity Index Measure
SVM	Support Vector Machines
TAL	Task Alignment Learning
TIR	Thermal infrared
TL	Transfer Learning
VAE	Variational Autoencoders
YOLO	You Only Look Once

Deep learning now overcomes these limitations through automation, precision, and real-time insights. By enhancing monitoring capabilities, improving disease detection, and optimizing production processes, modern AI technologies, DL contribute to more efficient and sustainable farming practices. As research continues to evolve, the potential to transform animal farming systems will pave the way for innovations that benefit both producers and consumers alike. In behavior analysis, deep learning models identify deviations in livestock movement patterns, signaling potential stress or environmental concerns. This capability allows farmers to implement proactive measures that improve animal welfare and operational efficiency (Akkem et al., 2024). Additionally, deep learning optimizes feeding practices by analyzing feed consumption and growth data. Refined feed formulations improve nutrition and reduce waste, leading to significant cost savings and increased profitability. Collectively, these innovations are essential for advancing sustainability and efficiency in livestock farming (Schillings et al., 2021). Accurate analysis is critical for making data-driven decisions and discovering the full potential of these data resources. Advanced techniques, like computer vision systems, are developing as powerful solutions for high-throughput phenotyping, which is required for optimal farm management and genetic improvement in breeding.

Images, initially captured for identification or monitoring, can yield additional contextual insights, significantly enhancing decision-making in animal behavior analysis (Abangan et al., 2023). For instance, images aimed at identifying individual animals and predicting their behavior often contain additional contextual data, such as housing conditions, seasonal indicators, weather conditions, animal stock density, and social network dynamics. By utilizing advanced technologies such as machine learning and computer vision, researchers can extract and analyze these insights, leading to improved management strategies and animal welfare outcomes (Guerri et al., 2024).

A new narrative synthesis review is required, in the domain of DL applications for animal farming is necessary due to the identified gaps and limitations in existing literature as describe in Research gaps in current literature (Section 2.5). The narrow focus limits comprehensive health and welfare assessments, as multimodal data integration could capture nuanced indicators like vocalizations for stress or IoT sensor data for physiological changes. This article address these gaps by exploring multimodal DL frameworks, covering a broader range of animal classes and applications, and promoting innovative approaches like transformers to enhance precision livestock farming.

This review systematically analyzes the role of deep learning in animal farming, focusing on key algorithms such as convolutional neural networks (CNNs), You Only Look Once (YOLO), memory-based neural networks (MBNNs), and generative adversarial networks (GANs). By examining 117 studies following PRISMA guidelines, this review identifies research trends, discusses significant advancements, highlights existing challenges, and explores future directions in deep learning applications for livestock farming. The article begins with an Introduction covering the role of technology in livestock and DL methodologies, including data collection and analysis. The Categorization of DL Algorithms section highlights key techniques like CNNs, YOLO, and GANs. The Applications in Animal Farming section is split into two parts: one focusing on practical uses such as detection, health monitoring, and behavior analysis, and another addressing challenges like environmental variability, data scarcity, and scalability. The Future Directions and Recommendations section suggests potential advancements and actionable insights, while the Conclusion summarizes the findings. This layout provides a comprehensive framework for exploring how DL can enhance efficiency and innovation in animal farming.

2. Methodology

To ensure methodological diligence, transparency, and consistency throughout the review process, systematic review was conducted in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines, which provide a comprehensive framework. The review aims to increase the certainty of its outcomes, provide a thorough and systematic approach to data collection and analysis, and establish a strong basis for drawing conclusions and directing future research by adhering to these criteria.

2.1. Data collection

A comprehensive search was carried out to find relevant information for this study across a variety of well-known scientific databases during 24/09/2024 to 30/09/2024. These databases include Google Scholar, ScienceDirect, Wiley, Scopus, Springer, MDPI, and IEEE as primary databases for this study was based on their credibility, interdisciplinary coverage, and relevance to deep learning applications in animal farming. These databases offers broad access to multidisciplinary research, provide high-impact journals in agriculture, cover animal science and computational methods, and ensures access to cutting-edge AI and computer vision studies. Together, these databases provide a comprehensive, peer-reviewed, and balanced foundation for this study, minimizing bias while maximizing scholarly rigor. The data in Fig. 1 reveal fluctuating trends in the number of articles published over year.

2.2. Search strategy and search strings

These queries were designed to capture both foundational research and recent advancements from 2020 to 2024. By spanning multiple databases and incorporating diverse terminologies, this strategy maximized the retrieval of literature relevant to various animal types and application domains.

A methodical filtering procedure and predetermined inclusion criteria were used to find and choose pertinent articles. To ensure thorough coverage of the literature, particular keywords were used in the search across several databases. Keywords "computer vision for animals", "deep learning algorithms for animal", "animal detection and recognition", "animal detection using DL algorithms", "deep learning for livestock", "image analysis for animal", "animal based computer vision", "animal diversity using deep learning algorithms", "animal classification using DL", as well as "deep learning applications in animal farming" were used. To ensure the inclusion of both foundational and contemporary achievements in the subject, these terms were inserted into each database's search engines to retrieve relevant papers.

A systematic search strategy was implemented using a set of carefully constructed search strings, tailored for each academic database. The objective was to ensure broad yet relevant coverage of studies addressing the intersection of deep learning (DL), computer vision (CV), and animal farming. Boolean operators and appropriate field tags were employed according to each platform's search protocol.

In Google scholar, the following query was used within the search field with Boolean operators (AND, OR) to ensure relevant results. "computer vision for animals" OR "deep learning algorithms for animal" OR "animal detection and recognition" OR "animal detection using DL algorithms" OR "deep learning for livestock" OR "image analysis for animal" OR "animal based computer vision" OR "animal diversity using deep learning algorithms" OR "animal classification using DL" OR "deep learning applications in animal farming" OR "CV in animal recognition and classification" OR "animal disease detection using DL" OR "animal counting" OR "CV in animal biodiversity".

In Web of Science, the search terms were included using the TS (Topic Search) field: "animal classification using DL" OR "deep learning applications "computer vision for animals" OR "deep learning for livestock" OR "deep learning algorithms for animal" OR "animal detection and recognition" OR "animal detection using DL algorithms" OR "image analysis for animal" OR "animal based computer vision" OR "animal diversity using deep learning algorithms" OR in animal farming" OR "Application of DL and CV in livestock recognition" OR "animal disease detection using image analysis" OR "animal counting" OR "computer vision in animal biodiversity". For IEEE Xplore, MDPI, Wiley, Springer the search was conducted across all metadata fields using the string: "computer vision for animals" OR "deep learning algorithms for animal" OR "animal detection and recognition" OR "deep learning for livestock" OR "animal classification using DL" OR

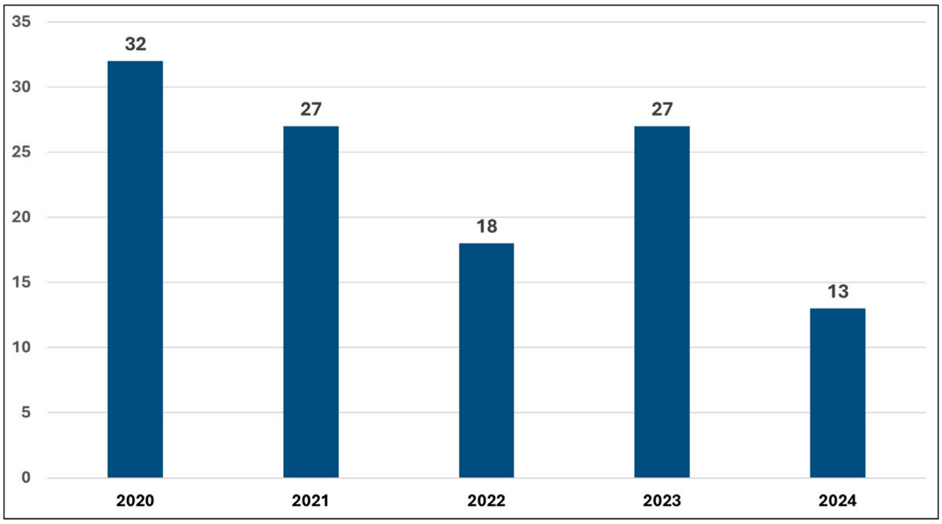

Fig. 1. Number of articles publish per year from (2020–2024).

Table 1
Inclusion and exclusion criteria for articles.

Articles exclusion criteria
Articles not irrelevant to livestock No English Language Review papers Not about farming Not related to DL algorithms Papers focusing on datasets. Articles published before 2020
Articles inclusion criteria
Papers relevant to livestock English Language only Related to DL algorithms Related to livestock farming May include data augmentation Articles from 2020 to 2024

"CV in animal biodiversity" OR "image analysis in animal farming systems".

The search was limited to english language papers published within the last five years (2020–2024) in order to guarantee the timely and relevant nature of the examined studies. A total of 173 articles were found through the database search, which were split as follows: 49 from Google Scholar, 47 from ScienceDirect, 15 from IEEE Xplore, 22 from MDPI, 21 from Wiley, and 19 from Springer. This approach ensured that the literature relevant to the application of DL algorithms in animal husbandry was comprehensive.

The exclusion criteria, as indicated in Table 1, were developed with careful consideration of the research topic, study objectives, and review scope. Using this method, 17 duplicate articles were found and eliminated. 13 papers were deleted after their whole contents were thoroughly reviewed, while 26 publications were removed following an evaluation of their abstracts and titles. After these exclusion criteria are applied, the remaining articles are determined to be eligible for systematic review or meta-analysis. At each step of the selection process, the number of publications that were found, screened, and included or eliminated is displayed using a PRISMA. At the time of writing, 117 research publications were still available in the literature following a thorough application of the inclusion and exclusion criteria. Fig. 2 shows final selection of publications with farming applications is displayed.

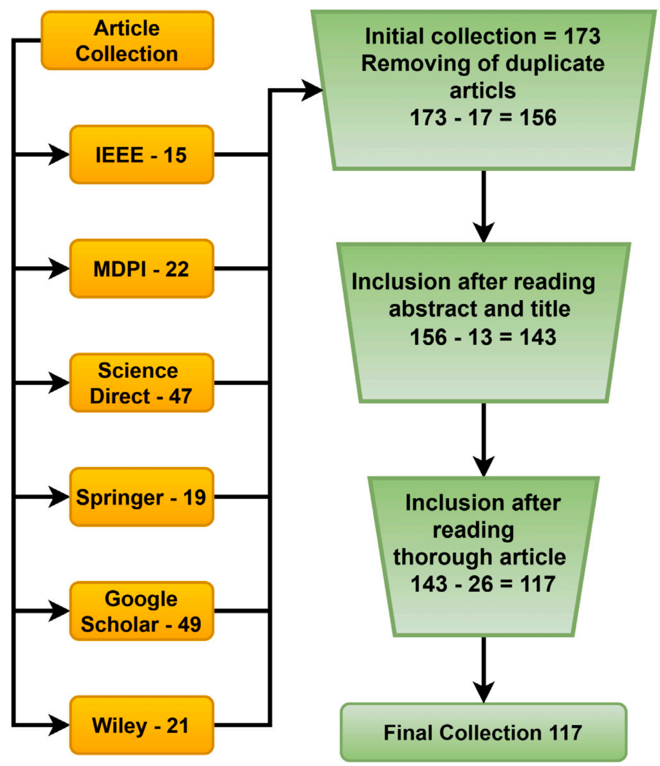

Fig. 2. Procedure for gathering a substantial number of articles.

2.3. Literature review

The literature show that DL has significantly advanced animal farming by automating tasks such as livestock detection, behavior monitoring, and health assessment, as noted in systematic reviews like (García et al., 2020; Liakos et al., 2018; Rohan et al., 2024). These studies predominantly utilize image and video datasets, leveraging convolutional neural networks (CNNs) for applications like cattle identification or anomaly detection in visual feeds (Palanisamy and Ratnarajah, 2021; Petso et al., 2022). However, audio data, such as vocalizations for stress detection, and sensory data from IoT devices, like accelerometers or temperature sensors, remain largely underexplored (Hossain et al., 2022; Zhang et al., 2024a). This heavy reliance on vision-based

Table 2
Review of earlier livestock-related DL framework work.

Reference	Application domain	Animal category	Audio	Video	Image	Sensors	Type
Valletta et al. (2017)	Animal behaviorists	Wildlife animals	✗	✔	✔	✔	Review
Wäldchen and Mäder (2018)	Animal identification	Agriculture and birds	✗	✗	✔	✗	Review
Liakos et al. (2018)	Agriculture & livestock production	Livestock	✗	✗	✔	✗	Review
García et al. (2020)	Grazing and animal health	Livestock	✗	✔	✗	✔	SLR
Ravoor and Sudarshan (2020)	Tracking animal movement	Wildlife animals	✗	✗	✔	✗	Review
Palanisamy and Ratnarajah (2021)	Detection of animals	Wildlife animals	✗	✗	✔	✗	SLR
Oliveira et al. (2021)	Biometric body measurements	Livestock	✗	✔	✔	✗	SLR
Petso et al. (2022)	Identification	Wildlife animals	✗	✗	✔	✗	Review
Hossain et al. (2022)	Precision livestock management	Livestock	✗	✔	✔	✔	SLR
Yousefi et al. (2022)	Detection and tracking	Livestock	✗	✔	✔	✗	SLR
Li et al. (2023)	Underwater biological detection	Aquatic animals	✔	✔	✔	✗	Review
Aguilar-Lazcano et al. (2023)	Animal Monitoring	Livestock and Pets	✗	✗	✗	✔	Review
Liu et al. (2024)	Localizing and tracking	Livestock and wildlife	✗	✔	✔	✗	Review
Rohan et al. (2024)	Behavior recognition	Livestock	✗	✔	✔	✔	SLR
Zhang et al. (2024a)	Animal Identification	Livestock and wildlife	✔	✔	✔	✔	SLR

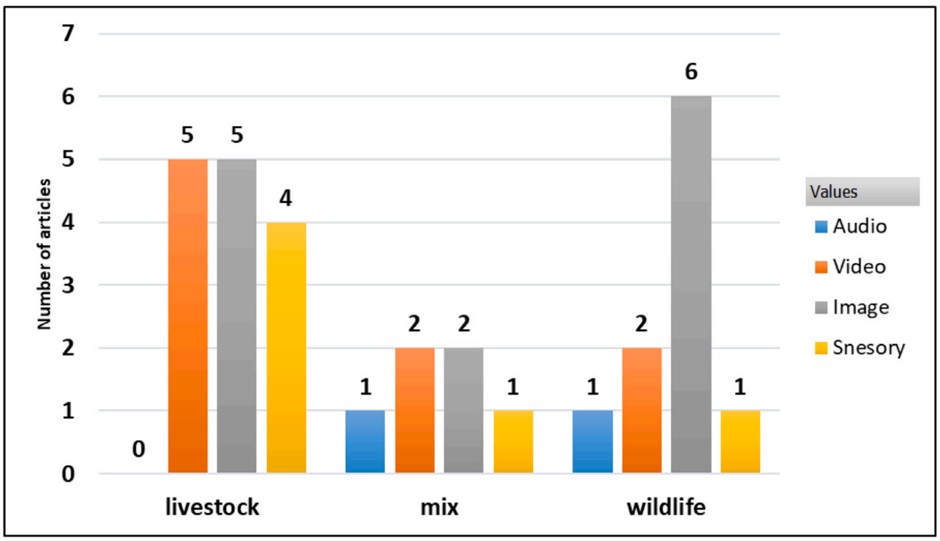

Fig. 3. Comprehensive dataset coverage across animal classes.

approaches limits the ability to capture nuanced indicators of animal health or behavior, particularly for early intervention in welfare issues.

The referenced Table 2 likely categorizes prior studies by application (detection, welfare, farm management), animal types (livestock, mixed, wildlife), and data modalities, highlighting the vision centric bias and minimal multimodal integration (Valletta et al., 2017; Oliveira et al., 2021; Liu et al., 2024; Aguilar-Lazcano et al., 2023; Yousefi et al., 2022). Research primarily focuses on commercially significant livestock (cattle, pigs, poultry), with limited attention to diverse species (Wäldchen and Mäder, 2018; Ravoor and Sudarshan, 2020; Zhang et al., 2024a; Li et al., 2023; Hossain et al., 2022). A broader range of animal classes and application contexts, as illustrated in Fig. 3, by advocating for multimodal DL frameworks, such as transformers (Taiwo et al., 2025), this work aims to enhance precision livestock farming and overcome the limitations of prior vision dominated research.

2.4. Analysis of collected article

Summarizing data helps provide a clear and concise overview of the trends, research focus, and gaps in the distribution of papers across different animal classes. Fig. 4 summarizes the distribution of articles focusing on various animal classes. The “multiple” category has the highest count of articles, indicating studies involving multiple animal classes. Among individual classes, chicken, fish, and cow dominate the research focus, with notable attention to pigs and cod. Less emphasis is placed on animals like rats, cats, rabbits, and pandas, each with minimal representation. This distribution highlights a certain farming animal classes receiving significantly more attention than others.

2.5. Research gaps in current literature

The above reviews on deep learning applications in animal farming have made considerable progress but often suffer from limitations in scope and coverage. Most existing works emphasize visual datasets focused on wildlife detection, species identification, and animal tracking, while overlooking other critical areas such as sensor-based monitoring, biometric analysis, and grazing or health management in livestock. A significant portion of the literature remains centered on image/video data and wildlife species, with limited integration of audio, physiological, and environmental sensor data on livestock. A clear gap exists in unified reviews that address multiple data modalities and span across all major animal systems. Key applications such as biometric body measurements, precision livestock management, animal health monitoring, underwater detection, and behavior recognition in farm settings are either underrepresented or fragmented across disciplines. This review bridges these gaps by:

- Including literature from 2020–2024, a period marked by accelerated AI adoption and integration with IoT, robotics, and edge computing.
- Categorizing applications across all animal domains and modalities, from underwater to livestock biometrics.
- Highlighting overlooked but critical applications like animal behaviorism, automated livestock monitoring, and cross-domain identification systems.

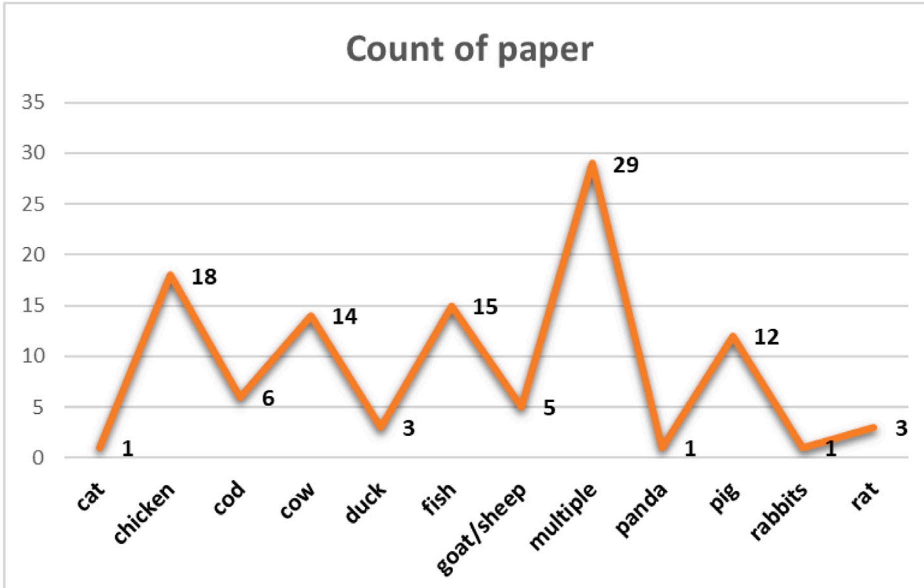

Fig. 4. Prominant animal classes discussed in collected articles.

- Offering a consolidated perspective valuable to researchers, agritech developers, and policymakers aiming to promote data-driven, ethical, and sustainable animal agriculture.

In essence, the fragmented and narrow scope of prior reviews necessitates this new, comprehensive study that not only synthesizes existing findings but also provides a roadmap for future innovations across species, sensor types, and smart farming applications.

3. Deep learning algorithms used in animal analysis

DL algorithms enable accurate analysis of large datasets for tasks like disease detection, behavior monitoring, and livestock tracking. By automating these processes, DL reduces manual effort, enhances decision-making, and addresses challenges like resource optimization and early problem detection in farming operations.

CNN family is the most dominant algorithms in animal farming, excelling in image-based tasks like disease detection and behavior analysis. YOLO is widely used for real-time object detection, while GAN is employed for data augmentation and synthetic dataset generation. These algorithms address critical challenges in accuracy and data limitations in animal farming. The numbers of DL algorithms applications in animal farming are shown in Fig. 5.

3.1. Advancements in CNN architectures

In the field of artificial intelligence, Convolutional Neural Networks (CNNs) have significantly transformed the processing of visual data, leading to advancements in image recognition, object detection, and video analysis. They utilize convolutional layers that apply filters to extract features, followed by pooling layers that reduce dimensionality while retaining essential information. This architecture enables CNNs to capture local patterns and spatial hierarchies within images, making them particularly effective for tasks where spatial relationships are critical.

3.1.1. Efficiency-orientated deep architectures

MobileNets are deep learning architectures designed for efficiency and accuracy. MobileNets use depthwise separable convolutions as shown in Fig. 6, which split the convolution operation into two parts (depthwise and pointwise), reducing computational cost and model size, making it ideal for mobile and embedded applications (Sandler et al., 2018; Howard et al., 2017). Xception enhances the Inception architecture by replacing traditional Inception modules with depthwise separable convolutions arranged in optimized blocks, leading to more efficient parameter use and improved performance (Chollet, 2017).

An EfficientNet model built using, uniformly scales depth, width, and resolution based on a neural architecture search, resulting in smaller models that achieve state-of-the-art accuracy (Tan and Le, 2019). Each architecture focuses on reducing computational cost while maintaining competitive performance, with MobileNets excelling in mobile devices, Xception offering advanced parameter efficiency, and EfficientNet optimizing scaling for high performance.

3.1.2. Complex CNN-based deep architectures

Complex CNN-based deep architectures refer to advanced designs that enhance performance, efficiency, and scalability for tasks such as image classification, object detection, and segmentation (LeCun et al., 1998). The structure of prominant Deep learning architecture shown in Fig. 7. AlexNet revolutionized CNN development with its success in the 2012 ILSVRC, showcasing the potential of deep learning for large-scale image classification (Krizhevsky et al., 2012). Its architecture, with five convolutional and three fully connected layers, introduced ReLU activation and dropout regularization, emphasizing the importance of GPU-powered training.

VGGNet is a deep CNN architecture using 3×3 convolutional filters and 2×2 max-pooling layers, with uniform depth across layers. It includes fully connected layers at the end, followed by a softmax classifier (Simonyan and Zisserman, 2014). VGG-16 and VGG-19 are common versions with 16 and 19 layers, respectively. ReLU activation functions are used after each convolutional layer. Known for its simplicity and strong performance on large datasets like ImageNet, VGGNet is effective in feature learning.

GoogLeNet is a deep CNN model that uses the Inception module, combining multiple filter sizes (1×1, 3×3, 5×5) and pooling operations in parallel for efficient feature extraction (Szegedy et al., 2014). It employs dimensionality reduction with 1×1 convolutions and global average pooling to reduce parameters. The network also includes auxiliary classifiers to improve gradient flow. Despite its depth, it maintains high accuracy and efficiency with fewer computational resources.

ResNet introduces residual connections, allowing input to skip layers and be added to the output, solving the vanishing gradient problem

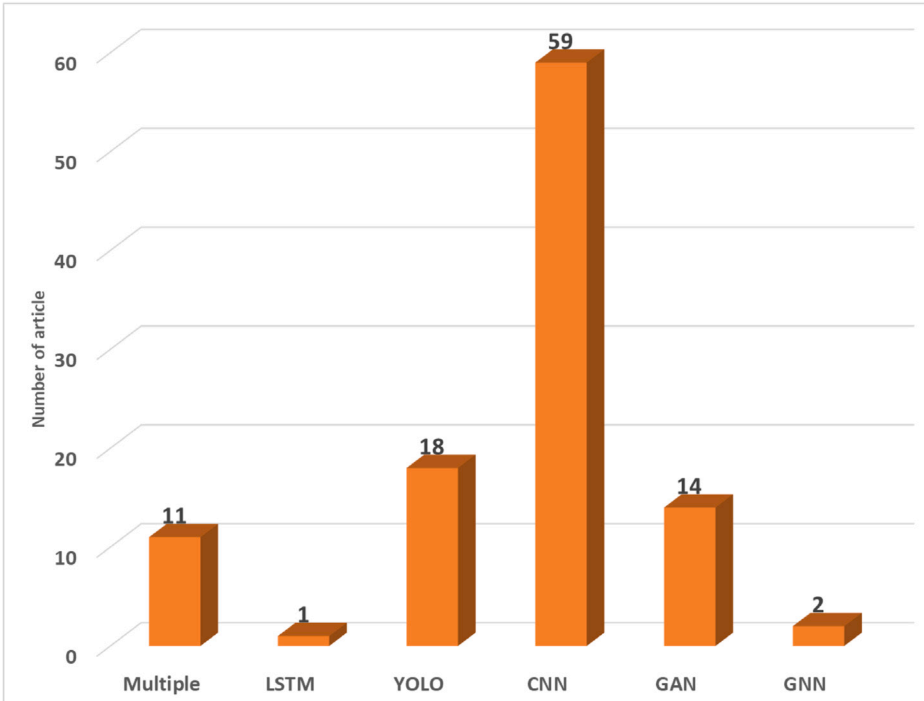

Fig. 5. Dominant deep learning algorithms used in animal farming.

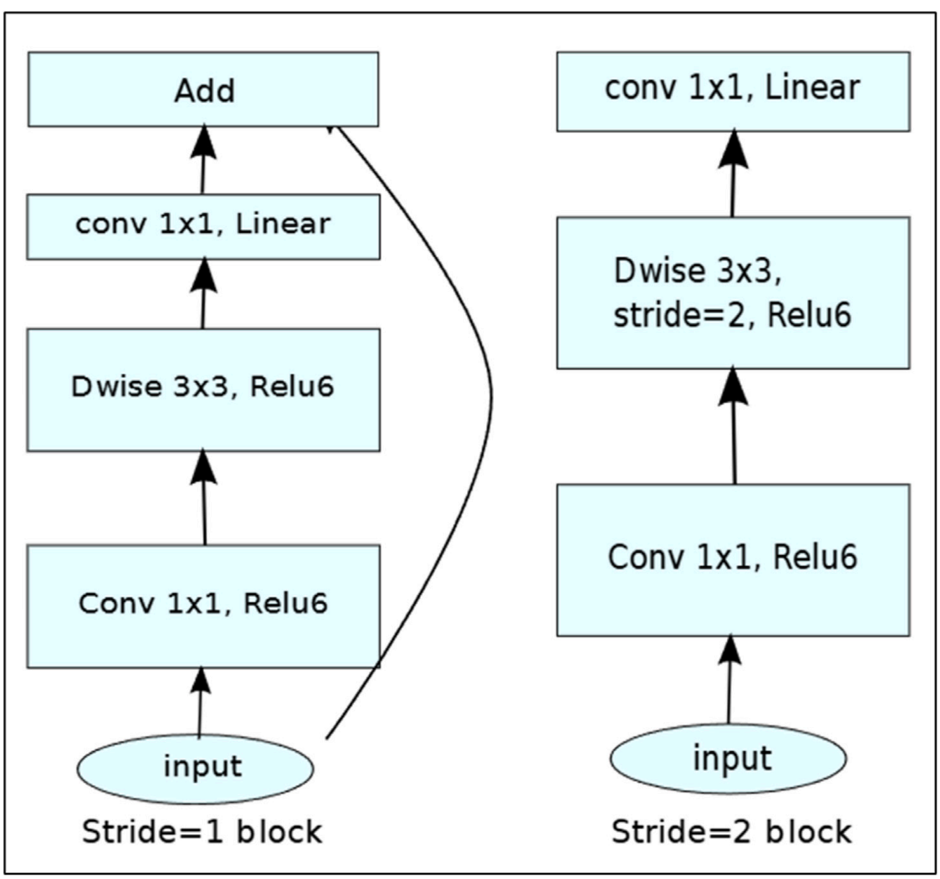

Fig. 6. The architecture of CNN based algorithm MobileNet (Sandler et al., 2018).

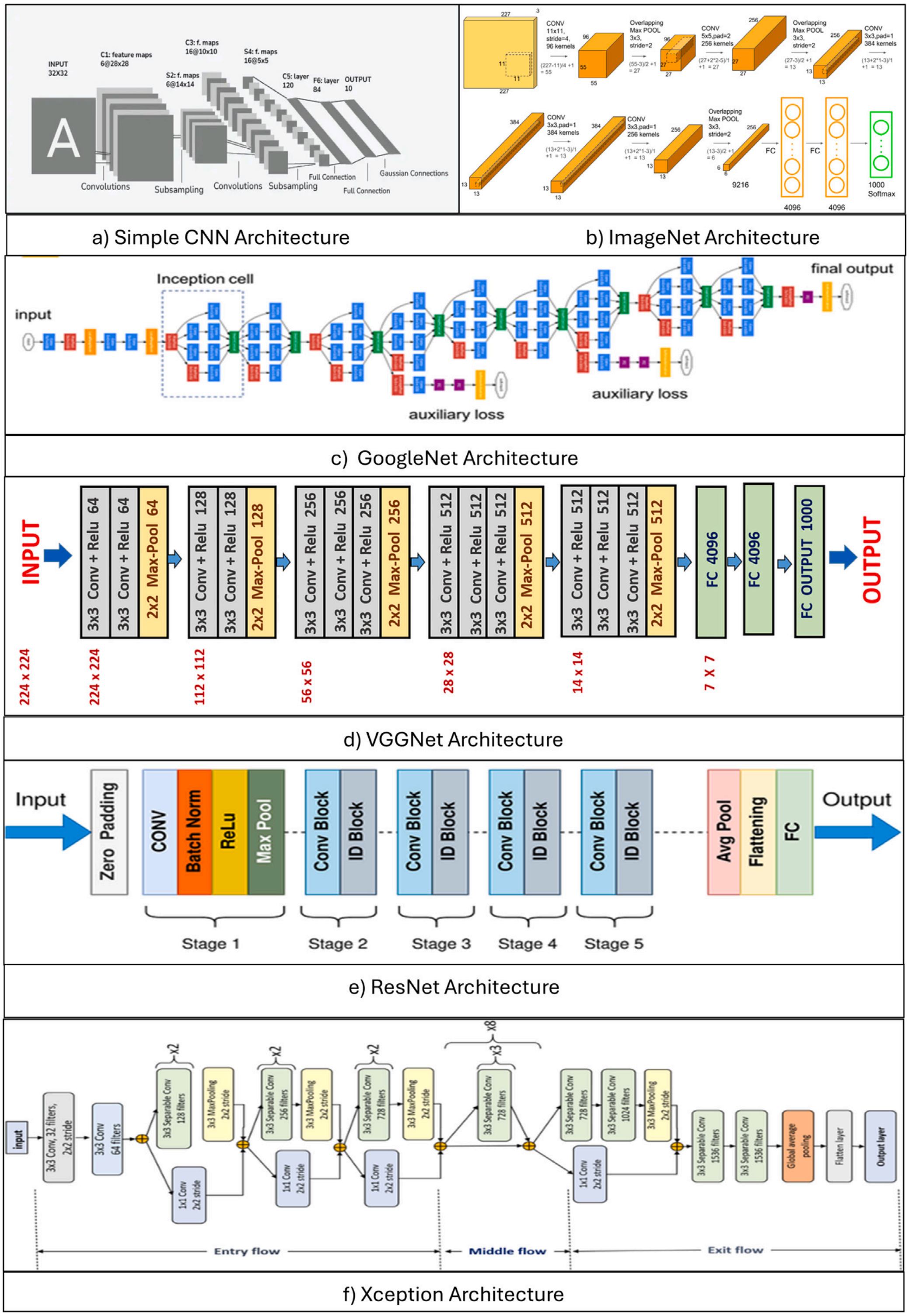

Fig. 7. Architecture of prominant CNN-based DL algorithms (a) Simple CNN (LeCun et al., 1998) (b) ImageNet (Krizhevsky et al., 2012) (c) GoogleNet (Szegedy et al., 2014) (d) VGGNet (Simonyan and Zisserman, 2014) (e) ResNet (He et al., 2015) (f) Xception (Chollet, 2017).

in deep networks. This enables the training of very deep networks without accuracy degradation (He et al., 2015). It uses bottleneck designs for computational efficiency in deeper models. ResNet achieved state-of-the-art performance in image classification tasks like ImageNet. The summary of these structures is briefly shown in Table 3. All these algorithm stated above designed for image classification applications.

The U-Net architecture is a convolutional neural network designed for image segmentation, characterized by its U-shaped structure with an encoder–decoder design and skip connections as shown in Fig. 8. The encoder extracts contextual features by downsampling the input image through convolutional and pooling layers, while the decoder reconstructs the output by upsampling and combining high-resolution features from the encoder via skip connections (Ronneberger et al., 2015). This allows U-Net to preserve spatial details and produce precise, pixel-wise segmentation maps. It excels in applications like biomedical imaging and performs well even on small datasets due to its

Table 3
Summary of prominant DL architectures used in animal farming.

Model	Author	Architecture details	Advantages	Disadvantages
Simple CNN	LeCun et al. (1998)	Few convolutional layers, max-pooling, and fully connected layers.	Easy to implement; works for small and large datasets.	Limited capacity for complex data; prone to overfitting on small datasets.
Alex Net	Krizhevsky et al. (2012)	8 layers (5 convolutional, 3 fully connected), ReLU activation, dropout, and overlapping pooling.	Revolutionized deep learning; introduced ReLU and dropout for faster training.	Computationally expensive; not efficient for mobile devices.
VGG Net	Simonyan and Zisserman (2014)	Deep network with 16–19 layers, small 3 × 3 convolution filters, and max-pooling layers.	Simple and effective; better feature representation with deep layers.	Very high computational cost and memory usage.
Google Net	Szegedy et al. (2014)	Inception modules with parallel convolutions of varying sizes; 22 layers deep.	Efficient with fewer parameters; good balance of speed and accuracy.	Complex architecture; harder to implement and modify.
Res Net	He et al. (2015)	Residual blocks with skip connections to address vanishing gradient problems; 50–152 layers.	Enables very deep networks; avoids vanishing gradients; better generalization.	Can still be computationally expensive; may overfit without large datasets.
U-Net	Ronneberger et al. (2015)	Encoder–decoder structure with skip connections for image segmentation.	Excellent for medical image segmentation; performs well on small datasets.	High memory usage; computationally intensive for high-resolution images.
Mobile Nets	Howard et al. (2017)	Lightweight architecture with depthwise separable convolutions; includes width and resolution multipliers.	Optimized for mobile and embedded devices; low computational cost.	Sacrifices accuracy for efficiency; limited for complex, high-resolution tasks.
Xception	Chollet (2017)	Deep separable convolutions and extreme Inception module with 36 convolutional layers.	Better accuracy than Inception and ResNet; efficient use of parameters.	Computationally demanding for certain tasks; not as lightweight as MobileNets.
Efficient Net	Tan and Le (2019)	Compound scaling of depth, width, and resolution; uses MBConv blocks.	State-of-the-art efficiency and accuracy trade-off; works well for mobile.	Scaling requires hyperparameter tuning; not ideal for very low-resource setups.

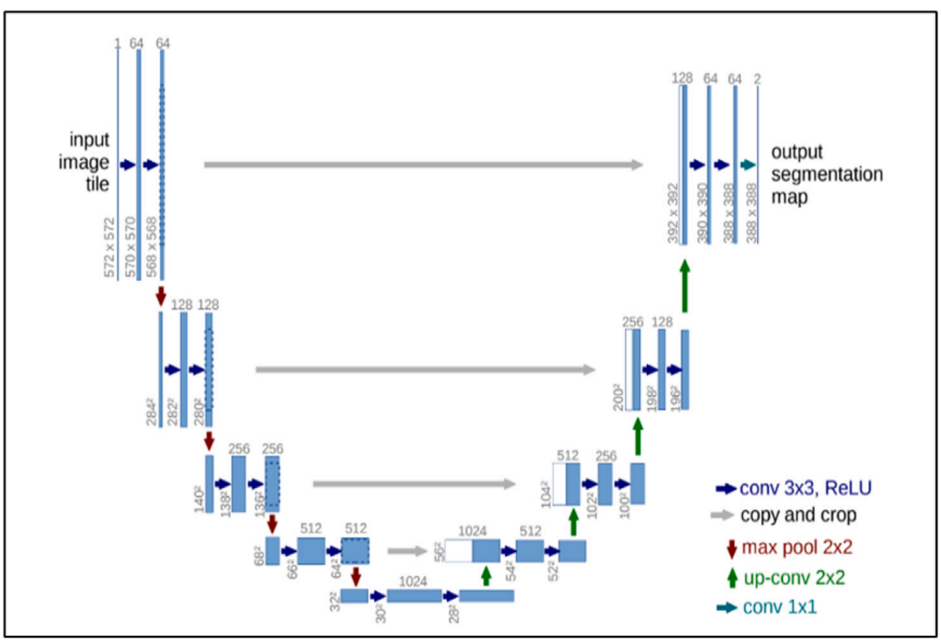

Fig. 8. U-Net a complex architecture of CNN (Ronneberger et al., 2015).

efficient use of feature maps and data augmentation. However, it is computationally and memory-intensive, especially for high-resolution inputs.

3.2. You Only Look Once

The YOLO (You Only Look Once) family of models has become a cornerstone in the field of computer vision, particularly in object detection. YOLO has continuously evolved, bringing improvements in speed, accuracy, and usability. Each iteration of YOLO has introduced new innovations, addressing limitations in previous versions and adapting to the needs of modern applications such as real-time detection, edge computing, and multi-task learning. Fig. 9 shows architectures of prominant YOLO variants.

The journey began with YOLOv1, introduced by Redmon et al. (2016). Unlike traditional object detection methods like R-CNN and Faster R-CNN, which relied on region proposal networks and multi-stage pipelines, YOLOv1 approached object detection as a single regression problem. The model divided the input image into a grid, with each

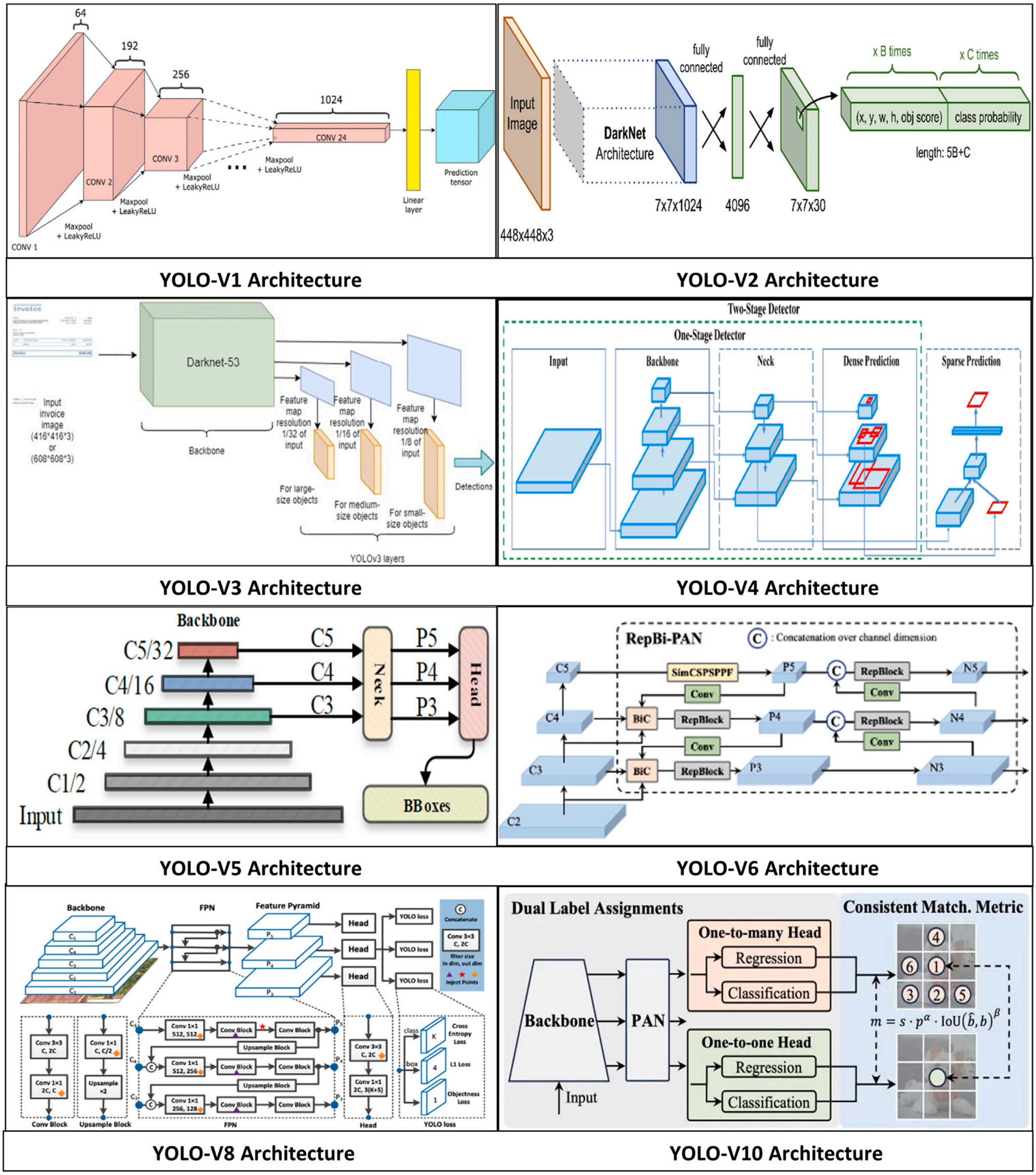

Fig. 9. Architectures variation of some YOLO framework (Redmon et al., 2016; Redmon and Farhadi, 2017, 2018; Bochkovskiy et al., 2020; Jocher et al., 2020; Li et al., 2022a; Varghese et al., 2024; Wang et al., 2024a).

cell predicting bounding boxes and class probabilities directly (Prabhu et al., 2023).

While YOLOv1 was revolutionary for its time, offering real-time performance, it struggled with detecting small objects and accurately localizing overlapping objects. These limitations stemmed from its coarse grid representation and lack of multi-scale feature extraction.

To address the shortcomings of YOLOv1, Redmon and Farhadi (2017, 2018) introduced YOLOv2, followed by YOLOv3. YOLOv2 introduced key improvements such as batch normalization (BN), anchor boxes, and a deeper backbone network, Darknet-19, which improved feature extraction. It also incorporated multi-scale training, enabling the model to adapt to images of different resolutions.

YOLOv3 further enhanced these capabilities by introducing a deeper backbone, Darknet-53, and integrating residual connections, which helped in mitigating the vanishing gradient problem. Most notably, YOLOv3 introduced multi-scale predictions, allowing the model to detect objects at varying sizes, significantly improving small object detection. These enhancements made YOLOv2 and YOLOv3 more versatile and accurate while maintaining their real-time performance (Pandiaraja et al., 2023).

Table 4
Prominant YOLO architectures used in animal farming.

Variant	Authors	Architecture	Advantages	Disadvantages
YOLO v1	Redmon et al. (2016)	Custom CNN with 24 convolutional layers and 2 fully connected layers.	Real-time object detection. Simplified architecture.	Struggles with small objects. Limited localization accuracy.
YOLO v2	Redmon and Farhadi (2017)	Darknet-19 backbone with 19 convolutional layers.	Improved accuracy with anchor boxes. Batch normalization reduces overfitting.	Challenges with very small objects.
YOLO v3	Redmon and Farhadi (2018)	Darknet-53 backbone with 53 convolutional layers and residual connections.	Multi-scale predictions enhance detection. Better handling of small objects.	Increased computational complexity.
YOLO v4	Bochkovskiy et al. (2020)	CSPDarknet-53 backbone with Spatial Pyramid Pooling (SPP) and Path Aggregation Network (PAN).	State-of-the-art accuracy. Enhanced feature fusion techniques.	Larger model size.
YOLO v5	Jocher et al. (2020)	CSPDarknet backbone with PANet and auto anchor box calculation.	User-friendly PyTorch implementation. Easy custom training and deployment.	Not officially from original YOLO authors.
YOLO v6	Li et al. (2022a)	EfficientRep backbone with decoupled head for classification and regression.	Optimized for edge devices. High inference speed.	Less focus on groundbreaking accuracy improvements.
YOLO v7	Li et al. (2022b)	E-ELAN backbone with Task Alignment Learning (TAL) and compound scaling.	Improved performance on small objects. Real-time detection capabilities.	Larger model size compared to YOLOv6.
YOLO v8	Varghese et al. (2024)	Advanced modular ConvNet supporting segmentation and detection tasks.	Highly modular and efficient. Supports multiple tasks.	Relatively new; long-term performance to be determined.
YOLO v9	Wang et al. (2024b)	Introduced Programmable Gradient Information (PGI) and Generalized Efficient Layer Aggregation Network (GELAN).	Enhanced efficiency and accuracy. Sets new benchmarks on datasets like MS COCO.	Increased architectural complexity.
YOLO v10	Wang et al. (2024a)	Introduced consistent dual assignments during training, eliminating the need for non-maximum suppression (NMS).	Improved real-time performance. Enhanced adaptability across various applications.	Specific disadvantages not well-documented.
YOLO v11	Khanam and Hussain (2024)	Incorporates C3K2 blocks, Spatial Pyramid Pooling Fast (SPPF), and advanced attention mechanisms like C2PSA.	Enhanced small object detection. Maintains real-time inference speed.	Increased architectural complexity.

YOLOv4 marked a significant leap in object detection (Bochkovskiy et al., 2020). YOLOv4 introduced a range of new techniques, including CSPDarknet53 as the backbone, spatial pyramid pooling (SPP) for better feature aggregation, and mosaic data augmentation, which improved the model's generalization to diverse datasets. Additionally, self-adversarial training (SAT) was introduced to enhance object localization. YOLOv4 achieved a remarkable balance between speed and accuracy, making it one of the most practical versions for real-world applications. It was designed with both research and deployment in mind, allowing for efficient training and inference on consumer-grade GPUs.

Although not officially part of the original YOLO lineage, YOLOv5 presented by Jocher et al. (2020), Madhumathi et al. (2023), became widely popular for its user-friendly framework and efficiency. YOLOv5 focused on scalability, offering multiple model sizes (e.g., YOLOv5s, YOLOv5m, YOLOv5l) to suit different resource constraints. Its lightweight architecture made it ideal for deployment on edge devices and production environments. However, YOLOv5 faced criticism for deviating from the original YOLO authors' work and for its lack of groundbreaking innovations compared to YOLOv4. Despite this, its practicality and ease of use cemented its place in the object detection community.

YOLOv6 (Li et al., 2022a) and YOLOv7 (Li et al., 2022b), pushed the boundaries of real-time object detection, focusing on lightweight architectures optimized for edge devices. These models introduced features like multi-task learning and improved small object detection, catering to applications requiring high-speed inference on limited hardware. YOLOv7, in particular, was lauded for being the fastest and most accurate real-time object detector at the time of its release. These versions were designed to maintain YOLO's trademark efficiency while incorporating advancements in training techniques and model design, making them suitable for a wide range of applications, from surveillance to autonomous systems.

The YOLO model architectures is discuss in Table 4 along with the author and the publication year. The model's architectural variations are explained, along with some of its benefits and drawbacks. The evolution of YOLO models presents significant opportunities for advancing monitoring systems. Building on YOLO-V7's efficient architecture, YOLO-V8 is expected to enhance context-aware detection in challenging farm environments through transformer based modules that better handle occlusions (e.g., animals behind fences or in crowded pens) and variable lighting conditions, while maintaining real-time performance on edge devices like drones or stationary cameras (Varghese et al., 2024; Kathir et al., 2024; Sharma et al., 2024b). This is particularly valuable for tracking individual animals in large herds or detecting subtle behavioral changes indicative of stress or illness. YOLO-V9 could further improve precision farming applications by integrating multi-task learning, combining detection with instance segmentation to monitor animal posture, feeding behavior, or lameness patterns over time critical for early disease intervention (Wang et al., 2024b). For energy-constrained deployments, such as battery-powered field sensors, YOLO-V10 may leverage lightweight transformer architectures (e.g., MobileViT) and sparsity-aware training to reduce computational costs without sacrificing accuracy in tasks like heat stress detection or automated weight estimation (Wang et al., 2024a). Looking ahead, YOLO-V11 could revolutionize livestock monitoring by adopting self-supervised learning to minimize reliance on manually annotated datasets, which are often scarce in agriculture. By incorporating spatial–temporal attention mechanisms, it could enable robust tracking of animals across long sequences, even in adverse weather or low-visibility conditions (Khanam and Hussain, 2024; Jegham et al., 2024; Guarnido-Lopez et al., 2024). Each iteration aims to balance speed, accuracy, and adaptability, with livestock-specific optimizations such as improved small-object detection for newborn animals or meta-learning for rapid adaptation across breeds making these models indispensable tools for modern precision livestock farming.

3.3. Autoencoder

Autoencoder (AE) is a type of artificial neural network used for unsupervised learning. Its primary goal is to learn a compact representation of input data and then reconstruct it as accurately as possible. The structure of autoencoders consists of two main components: an encoder and a decoder (Kingma et al., 2013). The encoder compresses the

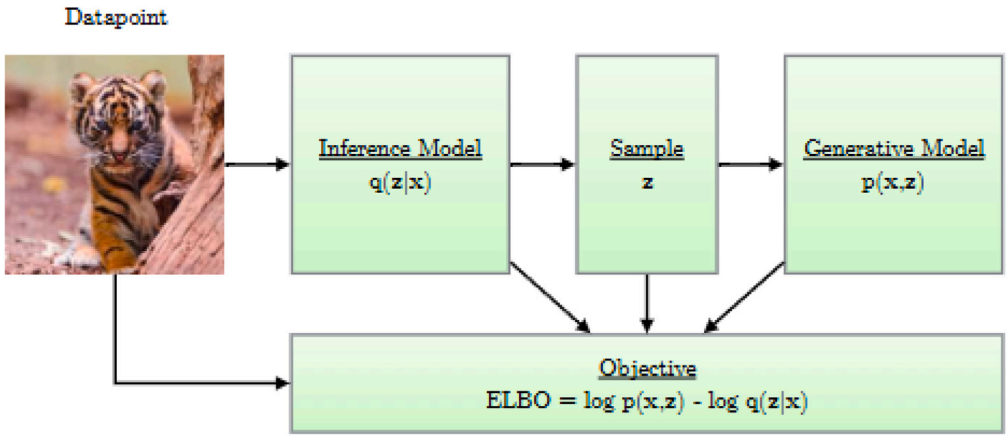

Fig. 10. Schematic of computational flow in a variational autoencoder (Kingma et al., 2019).

input data into a lower-dimensional latent space representation, capturing essential features while discarding redundancies. The decoder then reconstructs the original data from this compressed representation, aiming to minimize the reconstruction error between the input and output. In Denoising Autoencoders (DAEs), the structure remains similar but introduces noisy inputs during training. The encoder learns to map the noisy input into a latent space, and the decoder reconstructs the clean data, enhancing the model's robustness and feature extraction capabilities.

Variational Autoencoders (VAEs) extend this structure by introducing a probabilistic approach. Instead of mapping input data directly to a single latent representation, the encoder outputs parameters of a probability distribution (mean and variance) over the latent space (Kingma et al., 2019). Latent variables are sampled from this distribution, and the decoder reconstructs the data from these sampled variables as shown in Fig. 10. A VAE's loss function combines reconstruction loss with a regularization term (KL divergence) to ensure the latent space remains close to a prior distribution, often Gaussian. This structure enables VAEs to generate new, coherent data samples and interpolate smoothly between existing ones, making them highly effective for generative tasks. A Sparse Autoencoder (SAE) is a fully connected neural network that enforces sparsity in its hidden layer activations using techniques like L1 regularization or KL divergence, making it effective for feature extraction and anomaly detection. In contrast, a Convolutional Autoencoder (CAE) replaces fully connected layers with convolutional layers, preserving spatial information, making it ideal for image-based tasks such as denoising, super-resolution, and compression. While SAEs learn compact, sparse feature representations, CAEs leverage convolutional structures to extract hierarchical patterns, making them more efficient for visual data.

3.4. Generative Adversarial Networks (GAN)

Generative Adversarial Networks (GANs) consist of two main components: a generator and a discriminator, both implemented as neural networks. The generator maps random noise (sampled from a latent space) to data space, aiming to produce outputs that resemble real data. The discriminator, on the other hand, acts as a binary classifier, distinguishing between real samples from the dataset and synthetic samples generated by the generator (Goodfellow et al., 2014). During training, both networks engage in a minimax game: the generator aims to deceive the discriminator, while the discriminator strives to classify correctly. This iterative adversarial process pushes the generator to improve its outputs over time, resulting in realistic data generation when equilibrium is reached, Fig. 11 shown the some variants of GAN.

GAN variants adapt this fundamental structure for specific tasks. DCGANs use convolutional layers in both the generator and discriminator to handle spatial hierarchies in image data, improving the quality of

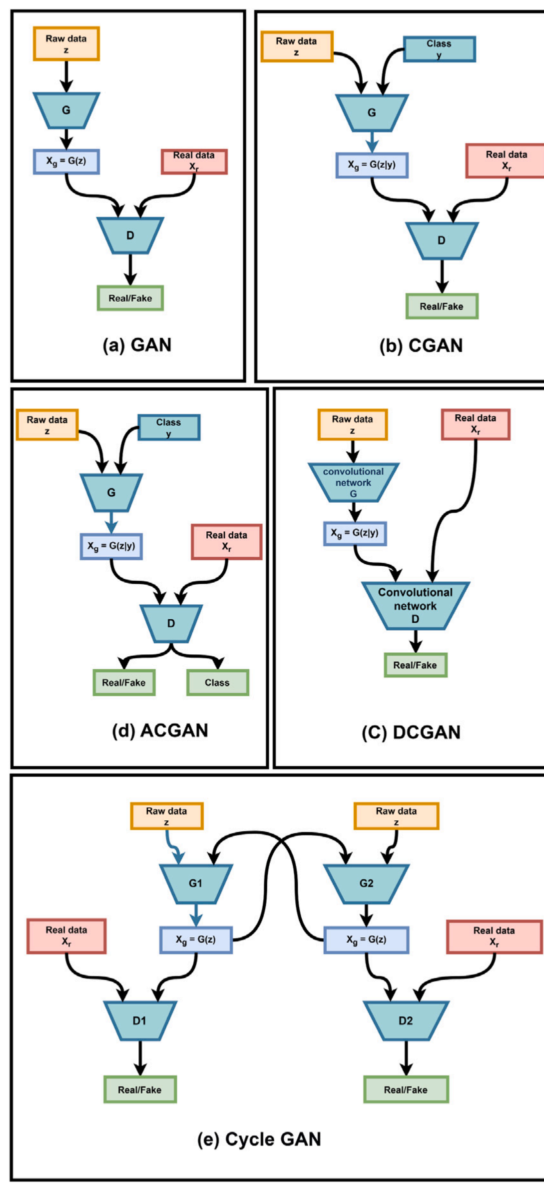

Fig. 11. Some well known GAN architectures (a) GAN (Goodfellow et al., 2014) (b) CGAN (Mirza and Osindero, 2014) (c) DCGAN (Odena et al., 2017) (d) ACGAN (Radford et al., 2015) (e) CycleGAN (Zhu et al., 2017).

generated visuals (Odena et al., 2017). CycleGAN introduces dual generators and discriminators to enable unpaired image-to-image translation, with cycle consistency loss ensuring transformations are reversible (Zhu et al., 2017). StyleGAN modifies the generator structure by adding a mapping network that transforms latent vectors into intermediate representations, enabling fine-grained control over image attributes (Mirza and Osindero, 2014; Radford et al., 2015). WGANs retain the basic GAN architecture but adjust the loss function to use Wasserstein distance, improving training stability and diversity of outputs. cGANs incorporate conditioning inputs (e.g., labels or text) into both the generator and discriminator, enabling targeted generation of samples based on specific criteria. These structural innovations have significantly enhanced GAN performance and broadened their application range.

3.5. Memory-Based Neural Networks (MBNNs)

Memory-Based Neural Networks (MBNNs) are a class of neural networks designed to integrate external memory mechanisms into traditional neural architectures, enabling them to handle tasks requiring long-term memory or complex reasoning. Recurrent Neural Networks (RNNs) are designed to process sequential data by maintaining a hidden state that evolves over time, enabling the capture of temporal dependencies. However, early RNNs were limited by the vanishing gradient problem, which hindered their ability to model long-range relationships (Bao et al., 2023). To overcome this, advanced architectures like Gated Recurrent Units (GRUs) and Long Short-Term Memory networks (LSTMs) introduced mechanisms to regulate information flow as shown in Fig. 12. LSTMs, in particular, use three gates; input, forget, and output gates to control what information is stored, discarded, or used, making them highly effective for sequence-based tasks such as machine translation and speech recognition.

Attention mechanisms further revolutionized sequential data modeling by enabling models to focus on specific parts of the input when making predictions. By dynamically assigning weights to input elements, attention allows models to prioritize relevant information while ignoring less important details. This has been particularly transformative in natural language processing (NLP) and image captioning tasks. Self-attention networks, a cornerstone of the Transformer architecture, go a step further by computing pairwise relationships between all sequence elements, capturing long-range dependencies efficiently and in parallel. This innovation has driven major advancements in NLP, powering state-of-the-art models for translation, summarization, and beyond.

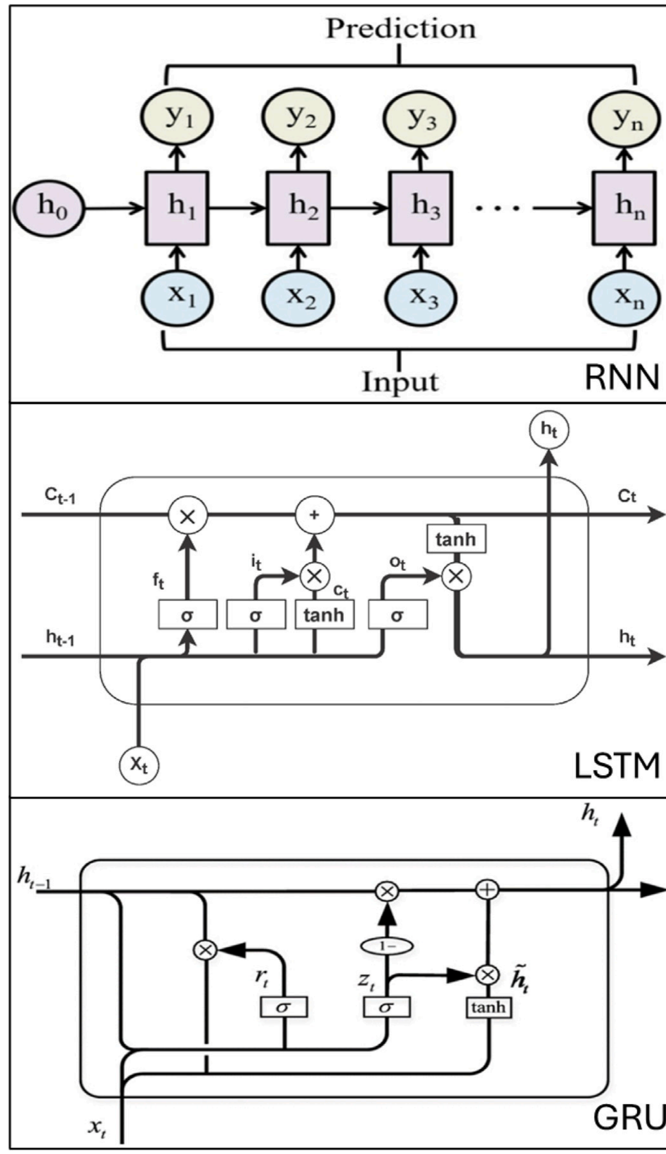

Fig. 12. Structure of some MBNNs (a) RNNs, (b) LSTMs, and (c) GRUs (Bao et al., 2023).

4. Application in animal farming

Deep learning (DL) is revolutionizing animal farming by enabling automated monitoring, early disease detection, and precision livestock management. DL models process data from sensors and cameras to track animal health, detect diseases, and analyze behavior in real time. Additionally, DL enhances livestock identification, breeding programs, genetic optimization, and environmental monitoring, leading to improved productivity and sustainability. This section explores various DL applications in animal farming, including livestock detection and identification using computer vision, health monitoring for disease detection and welfare assessment, generative models like GANs and autoencoders for synthetic data generation, and memory-based neural networks (MBNNs) for behavior prediction and long-term monitoring. Paragraph in all subsections of this section are arranged according to classification tasks, localization tasks and segmentation task, along with each animal class.

4.1. Livestock detection and identification

Accurate detection and identification of livestock are fundamental to modern precision agriculture systems. Deep learning (DL) models have been increasingly adopted to automate these tasks enabling real-time animal monitoring in diverse farming environments. An enhanced YOLOv4 model proposed by Ma et al. (2024) for automated pig face recognition to improve livestock management by addressing traditional methods' inefficiencies. Using a custom dataset, the model incorporates a lightweight design, data augmentation, and transfer learning (TL) for robust and accurate real-time processing. Research show high precision, recall, and F1 scores, making it suitable for edge-device deployment in agricultural settings. Seo et al. (2020) proposed a deep learning based pig detection system for embedded systems, offering a scalable and efficient livestock monitoring solution. The system uses a hybrid neural network architecture optimized for embedded systems, trained on diverse pig image datasets with robust data augmentation. Achieving over 95% detection accuracy and real-time inference, it offers a cost-effective and accurate solution for dynamic agricultural environments, enhancing animal management and welfare practices.

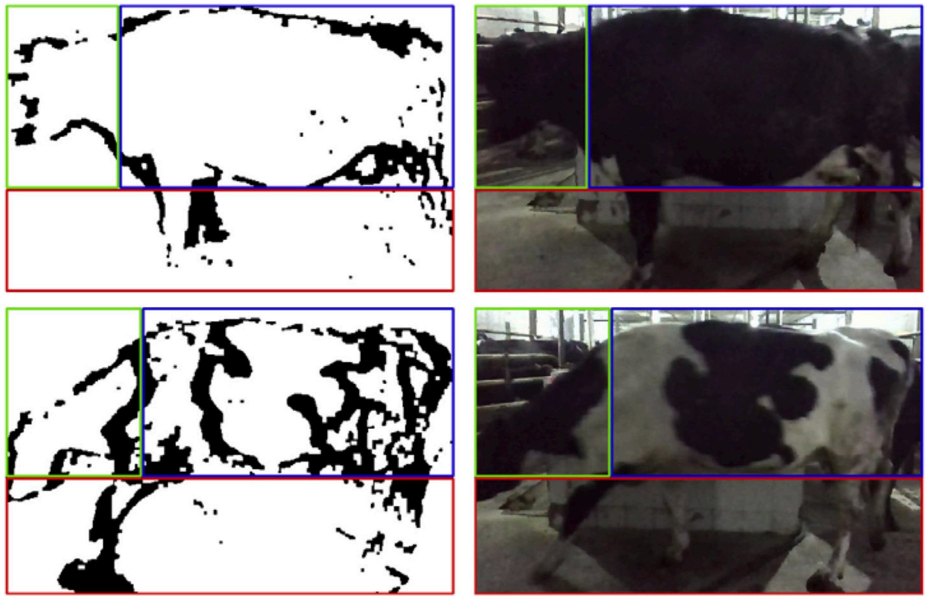

Fig. 13. Segmentation process for accurate cow identification (Hu et al., 2020).

In dairy farming, measuring individual cow feed intake is crucial for optimizing productivity and animal health. Traditional methods are labor intensive and error prone, but the use of RGB-D cameras and deep learning models offers a more efficient and accurate solution. One such system achieved a Mean Absolute Error (MAE) of 0.14 kg per meal in real world tests, demonstrating significant improvements over traditional methods and offering real-time monitoring capabilities for better feed management (Bezen et al., 2020). In dairy farming. Tassinari et al. (2021) introduced a computer vision system for real-time cow monitoring, achieving 95% accuracy and an F1 score of 0.92, with the system processing images at 30 FPS. This model outperformed traditional methods and provided actionable insights into cow behavior and health.

For effective herd management and health monitoring, deep learning-based automated identification systems have replaced traditional manual observation methods. These systems use CNNs to extract distinguishing features from cow images as shown in Fig. 13, achieving high identification accuracy even under challenging conditions. For example, combining VGG feature extraction with Support Vector Machine (SVM) classification has shown to yield impressive results, with feature fusion techniques further enhancing accuracy, particularly for cows with unique markings (Hu et al., 2020). GCT-VAE-GAN was introduced by Wang et al. (2023a) for low-light imaging in cattle farms, improving image quality with a dual-discriminator network and attention mechanism. The approach demonstrated robust performance in variable illumination conditions, significantly enhancing agricultural computer vision tasks.

Nasiri et al. (2024a) proposed a three step algorithm to estimate broilers' drinking time, using a U-Net architecture with ResNet50 for detection, Euclidean distance-based tracking, and a 3D convolutional network for behavior analysis. The algorithm achieved 83.21% accuracy in drinking time estimation and 88.32% in tracking individual drinking events. This approach offers a reliable, real-time tool for improving resource management and animal welfare on poultry farms. Elmessery et al. (2023) addressed poultry farming challenges with a YOLOv8-based model for broiler monitoring under varying light conditions. Trained on a dataset of 10,000 images and 50,000 annotations, the model achieved an mAP50 of 0.988 and F1 score of 0.972 for detecting broilers and classifying pathological conditions, demonstrating strong performance even in challenging environments with thermal and visual datasets.

The regulation of duck physiology through photoperiod management significantly impacts poultry farming. Duan et al. (2023) developed a multi-object tracking model based on YOLOv8 enhanced with the C2f-Faster-EMA module, achieving a mean average precision (mAP@50-95) of 97.9%. This model tracked duck activity by calculating travel distances and centroid positions, highlighting the impact of photoperiod regulation on meat duck welfare. Zheng et al. (2022) developed a deep learning-based method for automated sex classification of hemp ducks. Trained on a unique dataset of 17,090 whole duck images, the model, using YOLOv5 and VovNet_27slim, achieved an accuracy of 99.29%, an F1 score of 98.60%, and a processing speed of 269.68 frames per second. This method provides an efficient solution for accurate sex ratio estimation in farm environments, improving poultry farming operations.

BIRDSAI dataset was introduced by Bondi et al. (2020) to enhance wildlife monitoring using aerial thermal infrared (TIR) imagery. The dataset, collected via UAVs in African conservation areas, contains 48 annotated TIR video sequences with an average of 1,300 frames each. BIRDSAI addresses challenges such as motion blur, occlusion, and diverse animal behaviors, supporting the development of tracking algorithms. Results show its potential to advance automated surveillance systems for wildlife protection, offering a valuable resource for conservation efforts against poaching. Ferreira et al. (2020) used deep learning for small bird recognition, addressing species differences and environmental variability. By leveraging domain randomization, InceptionResNetV2, and data augmentation, their model achieved over 90% accuracy, supporting ecological studies and conservation. Kahl et al. (2021) presented BirdNET, a deep learning model for bird species identification from vocalizations as shown in Fig. 14, achieving 0.791 mean average precision for single-species recordings and 0.414 F0.5 score for soundscapes. Gupta et al. (2021) applied RCNNs for bird species classification, achieving 67% accuracy using hybrid CNN-RNN models, addressing background noise and class imbalance.

Machine learning's transformative potential was discussed by Tuia et al. (2022) for species detection, behavioral analysis, and predictive modeling of ecological changes. They also addressed the challenges of data scarcity, environmental variability, and the need for model interpretability in conservation contexts. Islam et al. (2023) utilized CNNs such as VGG16 and ResNet50 to automate species recognition in ecological studies, achieving a high accuracy of 94.2% on a diverse camera trap dataset. Despite this success, the challenge of distinguishing small-bodied animals like lizards from the background remained,

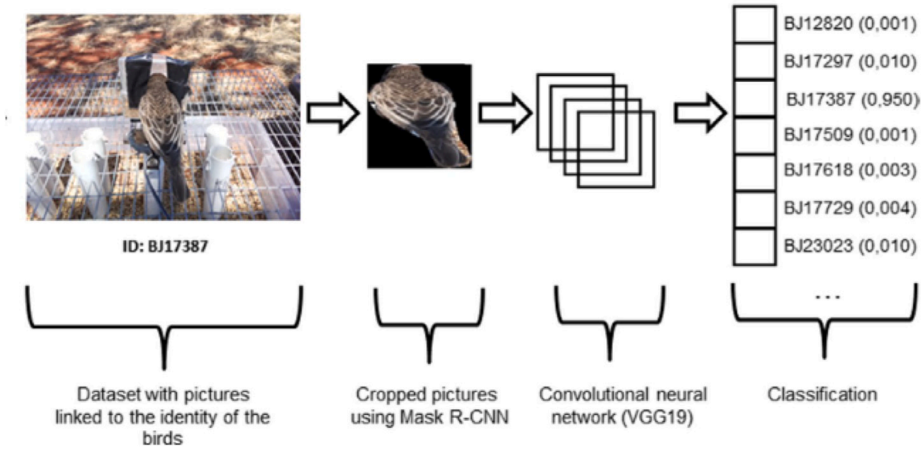

Fig. 14. Training convolutional neural networks to identify individual birds (Kahl et al., 2021).

with accuracy dropping to 83.1% due to size and color similarities. Ibraheam et al. (2020) highlighted the use of deep learning for animal species recognition in ecology and conservation, particularly for biodiversity monitoring through camera trap images. Despite challenges like imbalanced datasets and environmental variability, CNNs such as VGG16 and ResNet50 show promising results. With image preprocessing and augmentation, accuracies of up to 87% have been achieved, demonstrating deep learning's potential to address species recognition complexities and improve wildlife monitoring.

ThermalGAN framework was introduced by Khatri et al. (2022) to generate thermal images from FLIR videos, achieving 84.77% average precision and 94% F1-score in animal detection in thermal imagery, useful for nighttime monitoring. Gao et al. (2021) introduced GI-SleepNet, a GAN-assisted image-based sleep-staging program for mice, which generates synthetic REM sleep data to enhance classification accuracy. The program performs well even with data from a single mouse, offering a versatile solution for animal research contexts with improved accuracy. Clapham et al. (2020) developed BearID, a deep learning system for identifying individual brown bears through facial recognition. Trained on 4674 images of 132 bears, it achieved 0.98 precision in facial detection and 84% classification accuracy, enhancing wildlife monitoring and conservation efforts.

Goodwin et al. (2022) highlighted deep learning's growing role in marine ecology, helping with species detection, tracking, and environmental monitoring, despite challenges in data quality and computational demands. TRex was a real-time visual tracking tool used by Walter and Couzin (2021) for classification of 256 animals, using deep learning for accurate tracking in complex environments, applied across species like termites and fruit flies. Knausgård et al. (2022) used YOLO and CNN for fish detection, achieving 99.27% pre-training accuracy and 83.68%–87.74% post-training accuracy, overcoming challenges like lighting and species generalization. Besson et al. (2022) highlighted AI's role in automating species detection, tracking, and classification for biotic monitoring, offering standardized data for ecological studies. The summary of classification tasks is represented in Table 5 with applications.

To improve Animal Sound Classification (ASC) Kim et al. (2023) proposed a sound augmentation method using a class-conditional GAN, resulting in a 18.3% increase in accuracy and 13.4% improvement over traditional methods. Kopets et al. (2024) used StyleGAN2-ADA to generate synthetic sperm whale vocalizations, achieving a 2% deviation, which addressed the challenge of limited training data for marine biology. Corcoran et al. (2021) highlighted drones as an efficient wildlife monitoring tool, using deep learning (e.g., YOLO, Faster R-CNN) and image fusion for accurate species detection, offering real-time data with reduced costs.

Gunda et al. (2022) developed a deep learning-based framework for recognizing cows based on images of their muzzles and faces. The hybrid framework, which includes convolutional denoising autoencoders (CDAE), least squares GAN (LS-GAN), and Xception model for feature extraction, achieved an accuracy of 97.27% in cow recognition, outperforming other methods in insurance claim verification contexts. While for automatically milking system accurate teat detection is vital in rotary milking systems for effective milking (Lu et al., 2021). A deep learning based solution using CNNs has been proposed to automate teat detection, trained on a diverse dataset of cow images. The model achieves 95% precision and 93% recall, surpassing traditional methods, and operates in real-time to reduce human intervention and improve milking efficiency, benefiting both animal welfare and operational productivity.

A cattle identification system was proposed by Singh et al. (2021) using GAN for image enhancement and SIFT for feature extraction, achieving 98.86% identification accuracy with optimized processing time, outperforming existing methods. Similarly, Andrew et al. (2021) proposed a deep metric learning framework to identify individual cattle in large herds as structure shown in Fig. 15, achieving 97% accuracy and 94% precision in identifying Holstein-Friesian cattle without markers, highlighting the efficiency of deep learning in livestock management. Shang et al. (2023) proposed a deep learning-based method for individual cashmere goat identification, using SSD and transfer learning with Triplet Loss and Label Smoothing Cross-Entropy Loss. The model achieved a recognition accuracy of 93.75%. To enhance model feature learning, Cycle-GAN was used for low-accuracy instances. Mitra et al. (2022) developed a GAN-based anomaly detection model for hemorrhage detection in Doppler ultrasound images, achieving an Area Under the Curve (AUC) of 0.90 for immediate hemorrhage detection and an overall AUC of 0.83 across time points.

DL model for identifying individual animals was explored by Hou et al. (2020) to improve wildlife conservation and population management. Using a CNN tailored for panda identification, the model is trained on a curated panda image dataset with data augmentation for variations in posture, fur patterns, and lighting. Despite environmental challenges, the model demonstrates robust generalization and high identification accuracy, offering a scalable, non-invasive solution for wildlife monitoring and conservation. Camouflaged Object Detection (COD) tackles the challenge of identifying objects that blend into their surroundings, with applications ranging from wildlife monitoring to medical imaging. Despite advances in deep learning, challenges like data scarcity and real-time processing remain. Innovations like Chance-Constrained Game Net (CCGNet), utilizing modules such as Adaptive Feature Fusion (AFFM), Foreground Separation (FSM), and Edge Refinement (ERM), aim to improve detection accuracy, particularly in

Table 5
DL models used for animal detection and tracking in classifications tasks.

Author	Methodology/Model	Key results	Applications
Ma et al. (2024)	YOLOv4, lightweight design, data augmentation, transfer learning	High precision, recall, and F1 scores, suitable for edge-device deployment in agricultural settings.	Pig face recognition, livestock management
Tassinari et al. (2021)	Computer vision for real-time cow monitoring	95% accuracy, F1 score: 0.92, processing speed: 30 FPS, outperformed traditional methods.	Dairy farming, cow health monitoring
Hu et al. (2020)	VGG feature extraction + SVM for cow identification	High identification accuracy, effective for cattle with unique markings.	Livestock management, cattle identification
Nasiri et al. (2024a)	U-Net with ResNet50, 3D convolution for drinking time estimation	83.21% accuracy for drinking time, 88.32% for tracking drinking events	Poultry farming, resource management
Duan et al. (2023)	YOLOv8 with C2f-Faster-EMA for duck activeness tracking	mAP@50–95: 97.9%, tracked duck travel distances, linked to meat duck welfare	Poultry farming, welfare assessment
Elmessery et al. (2023)	YOLOv8 for broiler monitoring under varying light conditions	mAP50: 0.988, F1 score: 0.972, effective in thermal and visual datasets	Poultry farming, pathological condition classification
Zheng et al. (2022)	YOLOv5 + VovNet for hemp duck sex classification	99.29% accuracy, F1 score: 98.60%, processing speed: 269.68 FPS	Poultry farming, sex ratio estimation
Bondi et al. (2020)	BIRDSAI dataset, aerial TIR imagery for wildlife monitoring	Thermal infrared imagery for species detection, effective for automated surveillance systems.	Wildlife monitoring, anti-poaching
Ferreira et al. (2020)	Deep learning for small bird recognition with InceptionResNetV2 and data augmentation	Over 90% accuracy, handled species differences and environmental variability	Ecological studies, bird conservation
Kahl et al. (2021)	BirdNET, a deep learning model for bird vocalization analysis	0.791 mean average precision for single-species recordings	Eco-acoustics and biodiversity monitoring
Ibraheam et al. (2020)	CNNs (VGG16, ResNet50) with augmented data for animal species recognition	87% accuracy, improved model with data augmentation techniques	Ecological species monitoring
Khatri et al. (2022)	ThermalGAN for animal detection in thermal imagery (FLIR videos)	84.77% average precision, 94% F1 score for nighttime monitoring	Wildlife monitoring, thermal imagery
Gao et al. (2021)	GI-SleepNet (GAN-assisted sleep-staging for mice)	Improved accuracy with synthetic REM data, versatile for animal research contexts	Animal research, sleep-staging
Clapham et al. (2020)	BearID for brown bear facial recognition	0.98 precision, 84% classification accuracy	Wildlife monitoring, conservation efforts
Gupta et al. (2021)	RCNNs with hybrid CNN-RNN models	Achieved 67% accuracy across 100 species, with peak 90% for bird species.	Bird species identification and ecological surveys.
Goodwin et al. (2022)	Deep learning for species detection, tracking, and environmental monitoring	Addresses challenges in data quality and computational demands; enhances detection and tracking accuracy.	Marine species detection, tracking, and ecosystem monitoring.
Walter and Couzin (2021)	TRex for real-time visual tracking of up to 256 animals	Accurate tracking in complex environments; supports species like termites and fruit flies.	Real-time animal tracking across different species, aiding field studies.
Knausgård et al. (2022)	YOLO and CNN with Squeeze-and-Excitation	Achieved 99.27% pre-training accuracy and 83.68%–87.74% post-training accuracy despite challenges like lighting.	Fish detection in dynamic aquatic environments for ecological studies.
Besson et al. (2022)	AI for species detection, tracking, and classification	Automates species detection and classification, standardizing data for ecological studies.	Biotic monitoring and ecological data collection for research.

complex environments like surveillance and ecological studies Fan et al. (2020).

Darma et al. (2023) addressed the challenge of wildlife–vehicle collisions (WVCs) by proposing a machine learning-based animal detection system using CNNs, trained on diverse datasets. The system achieved a precision of 92%, recall of 88%, and an F1 score of 90%, demonstrating its effectiveness in detecting animals near roadways and minimizing false positives and negatives. This model holds promise for integration into autonomous vehicle technologies, improving road safety. Zhai et al. (2021) proposed a mutual graph learning framework for camouflaged object detection (COD), leveraging dual networks and graph neural networks (GNNs) to improve object–background differentiation. The approach significantly outperformed existing COD techniques, achieving superior accuracy and robustness in various scenarios, making it valuable for wildlife monitoring and surveillance applications.

CubeNet was introduced by Zhuge et al. (2022), an innovative method for COD, using an X-shape connection mechanism to better integrate multi-scale features for enhanced contextual understanding. CubeNet showed superior performance in detecting camouflaged objects, significantly improving detection metrics across various environments and outperforming existing COD models as shown in Fig. 16. Lv et al. (2021) propose a unified framework for camouflaged object detection (COD) that integrates localization, segmentation, and ranking. The framework improves object–background differentiation using advanced feature extraction and prioritizes objects based on visibility and relevance. It outperforms existing COD methods, showing superior accuracy and robustness across diverse datasets. Roy et al. (2023) introduced WilDect-YOLO, a specialized model for detecting species under threat in dynamic environments architecture. The model achieving a remarkable mean Average Precision (mAP) of 92.3%, an inference speed of 25 FPS, and a precision of 94.5% and recall of 90.1%.

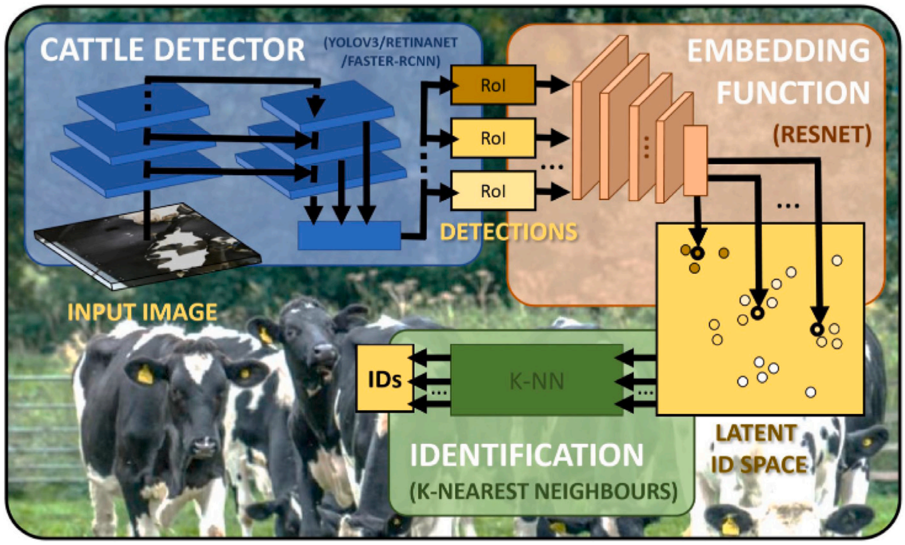

Fig. 15. Image clustering based on individual coat patterns is achieved via a dimensionality reduction algorithm (Andrew et al., 2021).

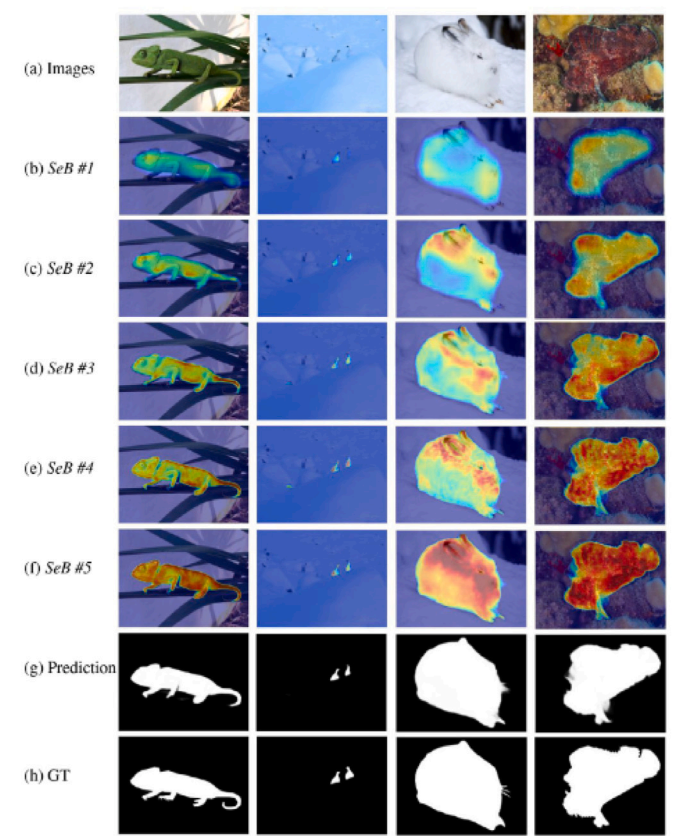

Fig. 16. Square Fusion Decoder (SFD) gradually highlights the items that are disguised and mutes noises. (a) Pictures; (b)–(f) Results in between taken from various Fusion Blocks; (g) Forecast; (h) The Ground Truth (Zhuge et al., 2022).

These results demonstrated the model's effectiveness for real-time, large-scale ecological monitoring.

The application of machine learning in autonomous underwater vehicles (AUVs) has significantly advanced the visual tracking of deep-water animals, offering new opportunities for ecological research and conservation. Machine vision technologies, combined with CNN and YOLO, enable efficient tracking of marine species in challenging environments, overcoming issues such as variable visibility, occlusion, and diverse animal behaviors (Katija et al., 2021). Integrating sensor fusion, which combines visual and acoustic data, has further enhanced detection accuracy and system robustness, contributing to more reliable monitoring in dynamic underwater conditions. Yari et al. (2024) proposed an automated real-time segmentation method for Atlantic salmon ovaries using ultrasound imaging, employing a modified U-Net enhanced with variational autoencoders (VAE) and GAN. The method achieved an average Dice score of 0.885 per individual, outperforming

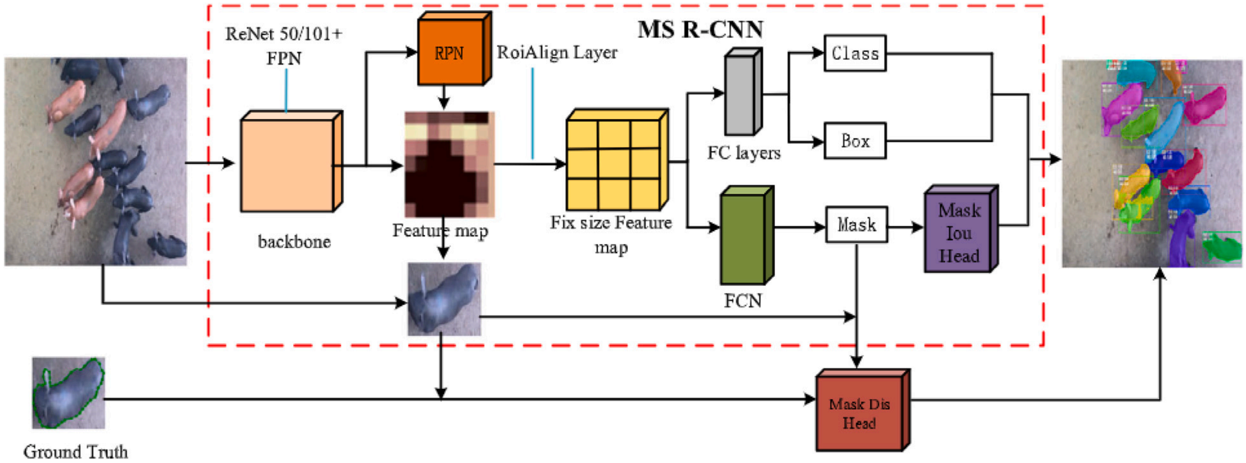

Fig. 17. Structure of MaskDis R-CNN (Tu et al., 2023).

traditional methods. This advancement offers an efficient, automated solution for monitoring salmon maturation in aquaculture, improving breeding management.

AISGAN was developed by Huang et al. (2023) for occlusion handling in piggeries, achieving a mean Intersection of Union (mIoU) of 0.823, outperforming traditional methods like Mask RCNN, with a 0.6% improvement in mIoU upon fine-tuning with unlabeled data. This approach improved occlusion-resistant spatial distribution analysis for animal monitoring. Yu et al. (2024) developed a deep learning-based cage-gate removal algorithm using Cyclic Consistent Migration Neural Network (CCMNN) to enhance image detection of combs and eyes in caged chickens. Applying the CCMNN network to YOLOv8 improved detection precision for combs (by 10.2%–12.8%) and eyes (by 2.4%–10.2%), contributing to better poultry health assessments. Tu et al. (2023) proposed an advanced instance segmentation method for pigs in complex scenes, using Mask Scoring Region-based Convolutional Neural Networks (MS R-CNN) combined with MaskDis, a generative adversarial network (GAN). This approach significantly improved segmentation quality, achieving a precision of 92.03%, recall of 92.18%, and F1 score of 0.9210, showing a 2.25% improvement in precision and 1.18% increase in F1 score over the basic MS R-CNN as shown in Fig. 17.

For chicken face detection Ma et al. (2022) introduced a GAN-MAE (Generative Adversarial Network-Masked Autoencoders) based object detection network, addressing dataset limitations and small object detection in smart poultry farming. The method, using the CSPDarknet53 backbone and an enhanced feature fusion module, achieved a mean average precision (mAP) of 0.84, a 29.2% improvement over other networks, with a detection speed of 37 frames per second. This approach is applicable in real-world poultry farming scenarios, offering significant improvements in detection accuracy and speed. Wu et al. (2023) introduced Super-resolution Chicken Detection for poultry detection in free-range farming. By using super-resolution fusion optimization and a GAN, the method enhanced image quality and improved detection accuracy. The validation showed a structural similarity of 99.9% and peak signal-to-noise ratios above 30, improving YOLOX models' precision with significant gains in Average Precision 95, making it a promising approach for multi-object poultry detection. The Table 6 reflects the DL applications in localization and segmentation tasks.

4.2. Health monitoring

DL algorithms have significantly improved animal health monitoring by enabling automated disease detection, and physiological condition assessment. DL assists in monitoring body temperature variations, while sound analysis models classify distress calls or respiratory issues and behavior analysis. The AI-driven approaches enhance real time health surveillance, reduce veterinary costs, and improve animal welfare in agriculture, conservation, and research settings.

A semi-supervised LSTM-Autoencoder algorithm was presented by Zhang et al. (2023) for early lameness detection in dairy cows using wearable sensor data. The system achieves 97.78% accuracy in detecting gait anomalies, offering an effective, real-time solution for continuous lameness monitoring, improving animal welfare and management. Wang et al. (2022) proposed using thermal imaging and deep learning to monitor ocular surface temperature (OST) in dairy cows, offering a non-invasive method to detect health issues like stress or infection. By employing CNNs to analyze thermal images under controlled conditions, the system outperforms traditional methods, enabling efficient and non-invasive health monitoring in dairy farming.

To detect early health issues (Alameer et al., 2020) introduced a deep learning-based system for tracking pig postures and drinking behaviors. The system achieves a high mean average precision (mAP) of 0.989, enabling scalable livestock health monitoring without additional sensors, improving animal welfare and farm management. Chen et al. (2020a) presented a CNN-based method for classifying pig behaviors (drinking vs. playful interactions) from video footage. The system overcomes challenges such as occlusion and environmental factors, offering an effective tool for monitoring animal welfare and early health issue detection in pigs.

An IoT-based predictive framework was introduced by Ahmod et al. (2021) for early disease detection in poultry using wearable sensors as shown in Fig. 18. By enhancing datasets with GAN-generated synthetic data, the system achieves 97% accuracy in classifying healthy vs. sick chickens, improving real-time disease detection. Zhang et al. (2024b) introduced an improved YOLOv8 model for detecting leg diseases in chickens, incorporating advanced techniques like Partial Convolution and Channel Prior Convolutional Attention, achieving a 7.2% increase in precision and operating at 66.8 fps for improved real-time performance. Biswas (2024) highlighted the importance of early detection of pet health issues and the challenges of processing large media datasets. They introduce the Media Insights Engine, which uses machine learning to streamline video and image analysis, improving diagnostic accuracy and speed while aligning with veterinary standards.

In red meat, Cui et al. (2024) proposed an autoencoder-assisted GAN (AE-GAN) for modeling polyunsaturated fatty acids (PUFAs). The AE-GAN enhances data generation, boosting R^2 values in regression models like Support Vector Regression, Random Forest, and Fully Convolutional Network by up to 0.2998, improving model performance in food science. Feighelstein et al. (2023) developed an automated system to detect acute postoperative pain in rabbits using video analysis. By addressing challenges like noisy footage and variability in pain expression, the system achieves over 87% accuracy, reaching 95% when using

Table 6
DL models for animal detection and tracking in segmentation and localization tasks.

Author	Methodology/Model	Key results	Applications
Lu et al. (2021)	Deep learning (CNN) for teat detection	Precision: 95%, recall: 93%, real-time operation, surpassed traditional methods.	Dairy farming, milking systems
Gunda et al. (2022)	Hybrid framework (CDAE + LS-GAN + Xception) for cow recognition	97.27% recognition accuracy, used for insurance claim verification	Livestock management, cattle identification
Singh et al. (2021)	GAN for image enhancement + SIFT for feature extraction	98.86% identification accuracy, optimized processing time	Livestock management, cattle identification
Andrew et al. (2021)	Deep metric learning framework, no markers for cattle identification	97% accuracy, 94% precision for Holstein-Friesian cattle	Livestock management, cattle identification
Shang et al. (2023)	SSD + Triplet Loss + Cycle-GAN for cashmere goat identification	93.75% recognition accuracy, enhanced learning for low-accuracy instances	Livestock management, goat identification
Mitra et al. (2022)	GAN-based anomaly detection for hemorrhage in Doppler ultrasound images	AUC: 0.90 for immediate detection, AUC: 0.83 across time points	Medical imaging, hemorrhage detection
Hou et al. (2020)	CNN tailored for panda identification	High accuracy, robust generalization despite environmental challenges.	Wildlife conservation, panda identification
Fan et al. (2020)	CCGNet with Adaptive Feature Fusion, Foreground Separation, Edge Refinement	Improved detection accuracy in complex environments, challenges with data scarcity.	Wildlife monitoring, medical imaging
Darma et al. (2023)	CNN-based detection system for wildlife–vehicle collisions	Precision: 92%, recall: 88%, F1 score: 90%, effective for autonomous vehicle integration for road safety.	wildlife–vehicle collision detection
Zhai et al. (2021)	Mutual graph learning framework for camouflaged object detection	Superior accuracy and robustness for wildlife monitoring and surveillance	Wildlife monitoring, camouflaged object detection
Zhuge et al. (2022)	CubeNet for camouflaged object detection	Improved multi-scale feature integration, enhanced contextual understanding	Camouflaged object detection
Lv et al. (2021)	Unified COD framework (localization, segmentation, ranking)	Outperformed existing COD methods, improved object–background differentiation	Wildlife surveillance, camouflaged object detection
Roy et al. (2023)	WilDect-YOLO (YOLO-based architecture)	mAP: 92.3%, precision: 94.5%, recall: 90.1%, inference speed: 25 FPS	Ecological monitoring, species detection
Katija et al. (2021)	Machine vision + CNN + YOLO for underwater species tracking	Overcame visibility and occlusion issues, integrated visual and acoustic data for improved accuracy.	Marine species monitoring
Yari et al. (2024)	Modified U-Net with VAE and GAN for salmon ovary segmentation	Dice score: 0.885 per individual, outperformed traditional methods	Aquaculture, salmon breeding management
Huang et al. (2023)	AISGAN for occlusion handling in piggeries	mIoU: 0.823, 0.6% improvement over Mask RCNN with fine-tuning on unlabeled data.	Animal monitoring, livestock management
Tu et al. (2023)	MS R-CNN + MaskDis GAN for pig segmentation in complex scenes	Precision: 92.03%, recall: 92.18%, F1 score: 0.9210	Pig segmentation, livestock management
Ma et al. (2022)	GAN-MAE-based object detection for chicken face detection	mAP: 0.84, 29.2% improvement over other networks, 37 FPS detection speed	Poultry farming, face detection
Wu et al. (2023)	Super-resolution chicken detection with GAN	99.9% structural similarity, improved YOLOX precision, significant gains in Average Precision50:95	Poultry detection, free-range farming
Yu et al. (2024)	CCMNN-based cage-gate removal algorithm for poultry health detection	SSIM: 91.14%, PSNR: 25.34 dB, improved detection of combs and eyes	Poultry health assessments

selected frames, offering a promising tool for pain assessment in rabbits. Tan and Liang (2023) proposed a framework combining transformer models and GANs for malaria diagnosis, achieving 99.8% accuracy with Swin Transformer and faster inference with MobileViT. These models offer high accuracy and resource efficiency for edge device deployment.

AnimalGAN was used by Chen et al. (2023b) a GAN model for simulating rat clinical pathology, providing an alternative to animal studies in drug safety evaluation. It outperforms conventional models in hepatotoxicity assessments, aligning results with human trends in a virtual experiment with 100,000 rats. Data-driven decision support systems, using machine learning and real-time monitoring, optimize livestock farming by improving health, welfare, and sustainability. Despite challenges like technical barriers and privacy concerns, these systems hold great potential for transforming agricultural practices by enhancing productivity and reducing waste (Niloofar et al., 2021).

Automated pig feeding behavior recognition, using deep learning techniques such as CNNs and YOLO, achieves high accuracy (99.4%) in real-time detection. This improves animal welfare by detecting health issues early, optimizes feed management, and contributes to sustainability in livestock farming (Chen et al., 2020b). Tail-biting detection in pigs, using CNNs combined with LSTM or image sequences, automates early identification of this welfare concern. A study achieved mean accuracies of 71.3% with CNN-LSTM, providing a simple and effective solution for farm implementation (Liu et al., 2020).

AI and computer vision was used (Nakrosis et al., 2023) to assess poultry health by classifying droppings into six categories based on abnormalities. With a classification accuracy of 91.78% using YOLOv5 and 92.41% in object detection, this system provides a robust solution for poultry health monitoring. Mao et al. (2022) presented a novel light-VGG11 CNN model for automatically identifying chicken distress calls as shown in Fig. 19, offering a faster and more efficient solution than traditional methods (Neethirajan, 2025). With 55.88% faster processing and comparable accuracy (95.07%), the model employs

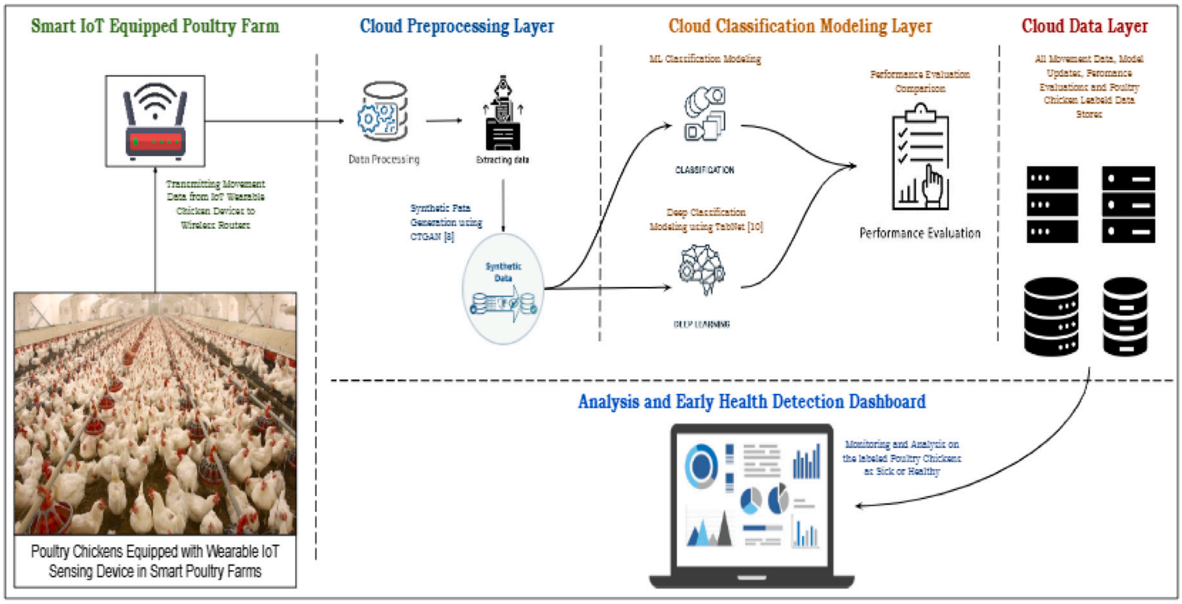

Fig. 18. The suggested Internet of Things-based predictive service to forecast the chickens' health in real time (Ahmod et al., 2021).

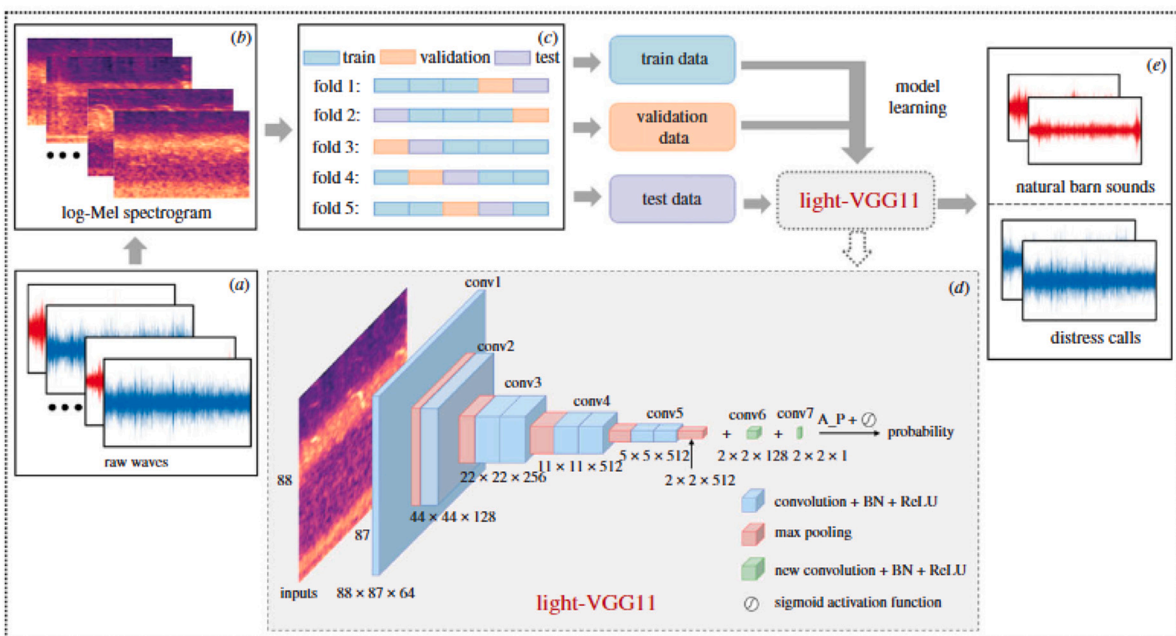

Fig. 19. The overall flow chart of the audio classification method. (a) Raw waveform signals. (b) Log-Mel spectrograms. (c) Fivefold cross-validation. (d) The architecture of light-VGG11. (e) Predictions (Mao et al., 2022).

data augmentation techniques, showing strong potential for monitoring large commercial flocks' welfare.

WGAN-GP model was developed by Hazra et al. (2022) to generate synthetic cat genome sequences with 93.7% correlation to original data. This model outperforms existing methods in sequence generation, offering a powerful tool for genetic research. Robson et al. (2021) applied a GAN-based computer vision pipeline to analyze ovine CT scans, achieving 98% similarity to manual analysis and reducing processing time from 30 min to 0.11 s, automating phenotype extraction for breeding programs. Li et al. (2024) estimated cow Body Condition Scores (BCS) using computer vision with a U-Net-based segmentation model, achieving high accuracy for measurements within 1–3 m. The approach offers effective near-range assessments, with potential for improvement at longer distances.

Using YOLOv7-tiny and SORT (Chen et al., 2023a) developed an automated system to monitor chicken movement detection and tracking structure is shown in Fig. 20, achieving 98.2% precision and reducing labor while enhancing farm management and pathogen risk mitigation. Degu and Simegn (2023) developed an automated poultry disease detection system using YOLO-V3 for ROI segmentation and ResNet50 for classification, achieving 87.48% precision and 98.7% accuracy, enabling fast and accurate disease detection for poultry farmers. Khanal et al. (2024) presented a transformer-based model for automated chicken counting in smart farms, overcoming challenges like lighting and occlusion. With a 96.9% accuracy rate, the model outperforms existing methods, offering a solution for precise chicken counting.

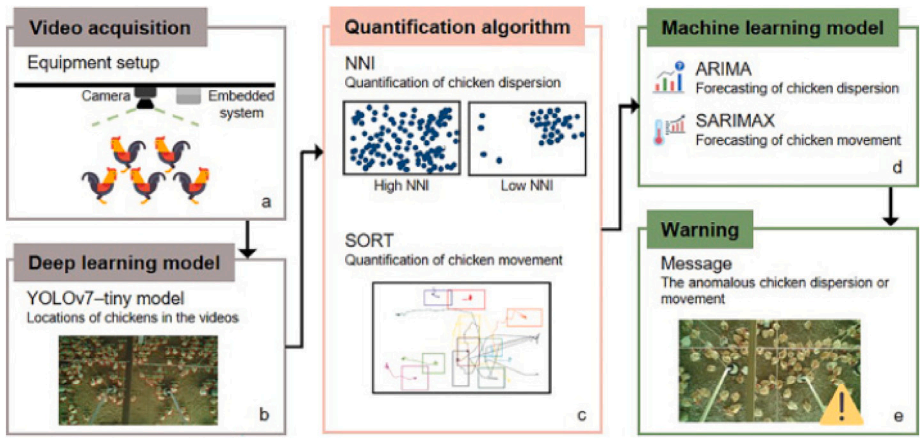

Fig. 20. Automatic chicken movement monitoring approach (Chen et al., 2023a).

Table 7
Deep learning application in animal health analysis.

Author	Focus area	Methodology	Results	Applications
Zhang et al. (2023)	Early lameness detection in dairy cows	Semi-supervised LSTM-Autoencoder on wearable sensor data	97.78% accuracy in detecting gait anomalies	Real-time monitoring for better animal welfare
Wang et al. (2022)	Monitoring ocular surface temperature (OST) in dairy cows	CNNs analyzing thermal images	Outperformed traditional methods in detecting stress and infections	Efficient, non-invasive health monitoring for dairy farming
Alameer et al. (2020)	Tracking pig postures and drinking behaviors to detect health issues	Deep learning system	Achieved mAP of 0.989	Scalable health monitoring in livestock
Chen et al. (2020a)	Classifying pig behaviors (e.g., drinking vs. playful interactions)	CNN-based video footage analysis	Effective under challenging conditions like occlusion	Animal welfare monitoring and early health detection
Ahmod et al. (2021)	Early disease detection in poultry	IoT-based predictive framework, GAN-enhanced datasets	97% accuracy in classifying healthy vs. sick chickens	Real-time poultry disease monitoring
Zhang et al. (2024b)	Detecting leg diseases in chickens	YOLOv8 model with Partial Convolution and Channel Prior Convolutional Attention	7.2% precision increase, 66.8 fps real-time performance	Enhanced real-time health management for poultry
Biswas (2024)	Early detection of pet health issues through media analysis	Media Insights Engine using machine learning	Improved diagnostic accuracy and speed	Streamlined veterinary diagnostics
Cui et al. (2024)	Modeling polyunsaturated fatty acids (PUFAs) in red meat	AE-GAN for enhanced data generation	Boosted R^2 values by up to 0.2998 in regression models	Improved performance in food science modeling
Feighelstein et al. (2023)	Detecting acute postoperative pain in rabbits	Automated video analysis	Over 87% accuracy, 95% accuracy with selected frames	Non-invasive pain assessment in rabbits
Tan and Liang (2023)	Malaria diagnosis with GANs	Transformer models combined with GANs	Achieved 99.8% accuracy using Swin Transformer	High-accuracy, resource-efficient malaria diagnostics
Chen et al. (2023b)	Simulating rat clinical pathology	GAN for virtual animal studies	Accurate hepatotoxicity assessments aligned with human trends	Replacing animal studies in drug safety evaluation

A semi-supervised hyperspectral imaging model using GAN and a 1D U-Net proposed by Campos et al. (2023) to detect foreign materials in poultry meat, achieving 100% precision and 96.9% balanced accuracy, with strong potential for food safety applications. Luo et al. (2023) developed a multi-source image-based method for detecting dead chickens on farms using deep learning. The method achieves 99.7% accuracy with NIR-depth images and provides valuable insights into optimal image sources for automated dead chicken detection. In the Table 7 a summary of Deep Learning application in animal health analysis is explain.

4.3. Animal diversity and counting

Deep learning (DL) algorithms have revolutionized the animal diversity and population estimation by enabling automated, accurate, and efficient analysis. It is widely used for identifying and classifying different animal species from images and videos. AI-powered systems integrated with drones, thermal imaging, and acoustic monitoring further improve species detection in diverse environments, aiding conservation efforts and ecological research.

Automated bird detection and counting highlighted by Akcay et al. (2020) using Faster R-CNN, it improving accuracy and efficiency over manual methods, while generating GIS-based distribution maps for large-scale bird surveys shown in Fig. 21. Norouzzadeh et al. (2021) proposed a deep active learning system for biodiversity monitoring, combining CNNs and active learning to optimize species identification and improve accuracy over traditional methods. Schneider et al. (2020) addressed challenges in automated species recognition from camera trap images, improving accuracy with advanced pre-processing and neural networks, highlighting the need for refinement in biodiversity monitoring tools. Xu et al. (2020a) developed a cattle counting system using Mask R-CNN and quadcopter imagery, achieving 96% classification accuracy and 92% counting accuracy. This system enhances livestock management, supporting precision agriculture and

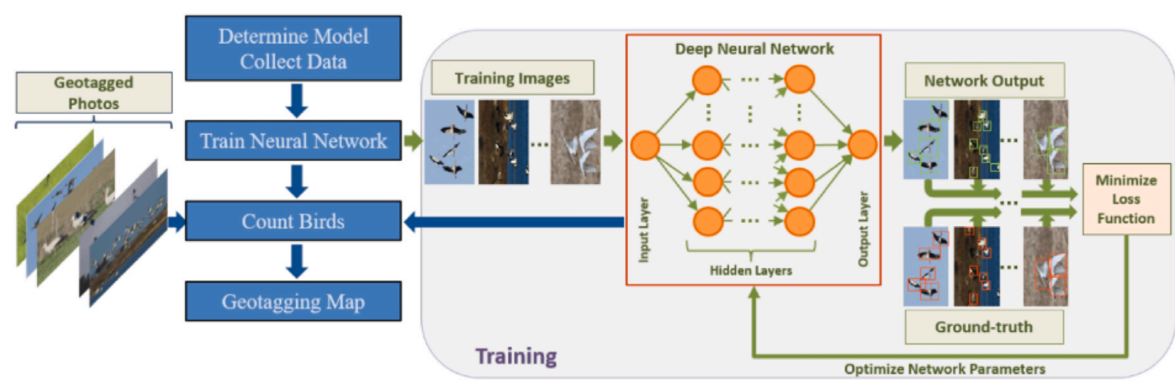

Fig. 21. Typical flowchart for the deep neural network training process in which the network adjusts its parameters (Akcay et al., 2020).

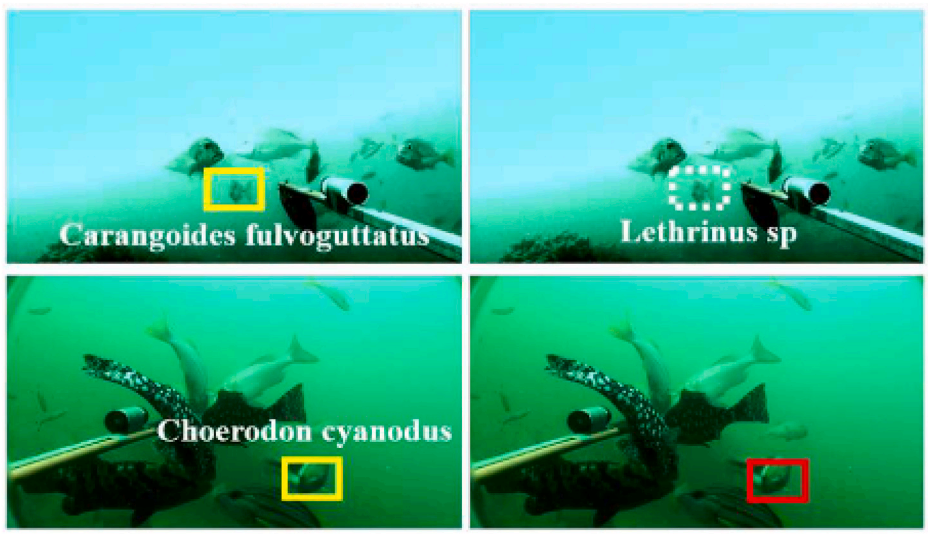

Fig. 22. Detect and classify fish species correctly within environmental variations (Jalal et al., 2020).

reducing labor costs. Xu et al. (2020b) used Mask R-CNN with quadcopters for livestock management, offering precise animal counting and classification in pasture environments.

High-resolution satellite imagery and CNNs was used by Duporge et al. (2021) to detect and count African elephants, achieving comparable accuracy to human observers, with robust generalization to lower-resolution imagery. Peng et al. (2020) improved kiang detection on the Tibetan Plateau using UAS imagery and a modified Faster R-CNN. The model outperforms traditional methods with high precision and recall, providing a reliable tool for wildlife monitoring and conservation.

CNN-based deep learning framework was proposed by Kandimalla et al. (2022) for automated fish population monitoring, offering efficient, accurate counting despite challenges like environmental variability. Martin-Abadal et al. (2020) introduced Jellytoring, an automated system using deep CNNs for jellyfish detection and quantification, achieving high accuracy (F1 score of 95.2%) for monitoring jellyfish populations. Freitas et al. (2023) introduced a deep learning system for automating pacu fish phenotyping, improving breeding efficiency through CNN-based body shape classification. While challenges like data variability remain, the system enhances trait selection decisions.

Hybrid neural network model was presented by Zhang et al. (2020) to estimate Atlantic salmon populations, addressing challenges like environmental variability and overlapping fish, offering improved precision and lower error rates. Han et al. (2020) proposed a deep CNN-based system for underwater organism detection, overcoming challenges like poor visibility and environmental variability. The system achieves 90% mean Average Precision and 50 FPS, integrated into an underwater robot for real-time marine conservation and research. Jalal et al. (2020) enhanced fish species detection with the YOLO framework and optical flow, improving accuracy by incorporating temporal data as shown in Fig. 22. This model outperforms traditional methods and aids fish population monitoring for conservation and aquaculture.

Automated fish abundance analysis was present by Ditria et al. (2020) using deep learning and object detection, surpassing human observers by 7.1% in single-image tests. The system improves accuracy and scalability for fish population monitoring in aquatic ecology. Zhang et al. (2021) introduced a CNN method for shrimp egg counting in hatcheries, using density map regression to handle occlusions. The Shrimp Egg Counting Network (SECNet) achieves 99.2% counting accuracy and integrates a low-cost, real-time setup, offering an efficient solution for aquaculture egg counting. The Table 8 shows the summary of section Deep learning applications in animal diversity and counting.

4.4. Behavior and pose estimation

AI revolutionized the analysis of animal behavior and pose estimation by providing accurate and automated tracking of movement, posture, and interactions. Pose estimation models leverage deep neural networks to track skeletal movements, aiding in gait analysis, stress detection, and activity monitoring. These techniques are crucial for applications such as livestock welfare assessment, and wildlife conservation.

Table 8
Deep learning applications in animal Counting and biodiversity monitoring.

Author	Focus area	Methodology	Results	Applications
Akcay et al. (2020)	Automated bird detection and counting	Faster R-CNN for bird detection	Improved accuracy and efficiency over manual methods	Large-scale bird survey and GIS mapping
Xu et al. (2020a)	Livestock counting and management	Mask R-CNN with quadcopter imagery for livestock counting and classification	Achieved 96% classification accuracy and 92% counting accuracy; enhanced livestock management.	Precision livestock management, reducing labor costs and improving efficiency.
Duporge et al. (2021)	African elephant detection using satellite imagery	CNNs applied to high-resolution imagery	Comparable accuracy to human observers	Wildlife conservation and biodiversity monitoring
Schneider et al. (2020)	Species recognition from camera traps	Advanced pre-processing and neural networks for image analysis	Improved accuracy in species recognition from camera trap images.	Biodiversity monitoring through automated image classification.
Peng et al. (2020)	Kiang detection on Tibetan Plateau	Modified Faster R-CNN with UAS imagery	High precision and recall, outperforming traditional methods.	Wildlife monitoring and conservation in high-altitude regions.
Norouzzadeh et al. (2021)	Biodiversity monitoring	CNNs and active learning combined for optimizing species identification	Improved accuracy of species identification over traditional methods.	Biodiversity monitoring with enhanced species recognition.
Kandimalla et al. (2022)	Fish population monitoring	CNN-based deep learning framework	Efficient and accurate counting despite environmental variability	Aquatic ecology and fish population conservation
Martin-Abadal et al. (2020)	Jellyfish population monitoring	Deep CNNs for automated jellyfish detection and quantification	95.2% F1 score	Real-time marine conservation and research
Freitas et al. (2023)	Automating pacu fish phenotyping	Deep CNN for body shape classification	Improves breeding efficiency; challenges include data variability.	Enhances trait selection decisions for breeding programs.
Zhang et al. (2020)	Estimating salmon populations	Hybrid neural network model for population estimation	Improved precision and lower error rates, overcoming environmental variability.	Population estimation in salmon farming and conservation.
Han et al. (2020)	Underwater organism detection	Deep CNN with 90% mean Average Precision and 50 FPS for real-time detection	Real-time performance integrated into underwater robots; achieved 90% mAP.	Marine organism detection and conservation efforts using underwater robots.
Jalal et al. (2020)	Fish species detection	YOLO framework with optical flow for fish detection	Incorporating temporal data improved accuracy, surpassing traditional methods.	Fish population monitoring in aquaculture and conservation settings.

Fig. 23. A few qualitative findings from the Animal Pose dataset training HRNet-w32 (Yu et al., 2021).

The AP-10K dataset was developed by Yu et al. (2021), with 10,015 images from 23 families, advanced animal pose estimation shown in Fig. 23, providing a foundation for multi-species training and human pose model application. Lauer et al. (2022) enhanced DeepLabCut for multi-animal posed estimation, addressing occlusions and appearance similarity. Their method improved tracking accuracy in animal behavior studies the results shown in Fig. 24. With ML and CV techniques, including semantic segmentation and regression models (e.g., ResNet-101-D), Wang et al. (2021) provided non-invasive solutions for accurate livestock weight prediction, with errors as low as 13.11 pounds. Li and Lee (2021) proposed a UDA approach with a Multi-Scale Domain Adaptation Module (MDAM) and network to improve animal pose estimation, achieving better accuracy and generalization across species on the TigDog and VisDA 2019 datasets.

A non-invasive CNN method was proposed by Meckbach et al. (2021) for livestock weight estimation, achieving high accuracy with reduced animal stress, addressing challenges like data imbalance and environmental variability. Deep learning models, leveraging CNNs and RNNs, automate livestock behavior recognition, offering superior performance in diverse environments and real-time applications for livestock welfare and health monitoring (Chen et al., 2021). SLEAP developed by Pereira et al. (2022), a multi-animal pose tracking framework that achieved high accuracy, with less than 2.8 pixels of error on 95% of points, and processed images at speeds up to 320 FPS. These quantitative results highlight how machine learning is revolutionizing ecological monitoring and advancing our understanding of animal behavior in social and environmental contexts.

CNN-based system for dairy cow identification and feeding behavior monitoring introduce by Achour et al. (2020), achieving 97.95% accuracy with fast training (6 s) and low computational cost (6.51 MB), improving real-time livestock management. Fuentes et al. (2020) utilized spatio-temporal information and CNNs to recognize cattle behaviors, achieving over 95% precision. This system, with data augmentation, enhances real-time monitoring for health management in farming. Manoharan (2020) explored Res-DenseYOLO for dairy cow behavior recognition, achieving 94.7% precision and 96.3% mAP, enhancing health monitoring with over 98% accuracy in feeding behavior detection. Jiang et al. (2020) highlighted deep learning for goat behavior recognition, using IMUs for movement tracking and YOLOv5 for counting and tracking, achieving a 92.19% mAP and providing real-time insights into health and behavior.

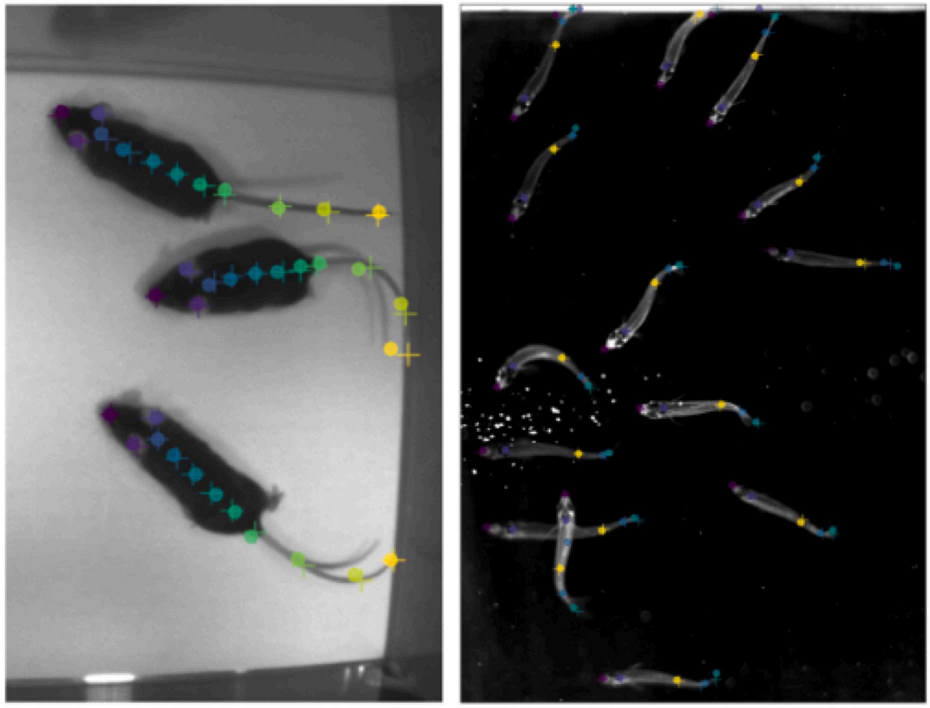

Fig. 24. DeepLabCut for finding multi-animal posed estimation (Lauer et al., 2022).

Noor et al. (2020) demonstrated the use of CNNs for sheep facial expression classification structure shown in Fig. 25, achieving over 90% accuracy in detecting emotional states like stress and happiness, improving animal welfare and farm management. A deep learning system for broiler chicken behavior monitoring, utilizing CNNs and data augmentation, improves pose estimation and real-time performance, enhancing poultry welfare management (Fang et al., 2021). An automated system was developed by Nasiri et al. (2024b) for quantifying broiler behavior, achieving 96.7% accuracy and 92.4% AUC for monitoring welfare behaviors such as stretching and preening.

A deep learning approach using a lightweight CNN with attention mechanism was introduced by Rickert et al. (2020) for automatic pig posture detection, achieving 97.7% accuracy for standing pigs and 84.1% for mounting pigs. With an inference speed of 63 ms, this system enables continuous monitoring of pig behavior and welfare. The handle challenges, such as methodological inconsistencies and limited access to computational resources, Von Ziegler et al. (2021) reported that 80% of research labs lack standardized protocols, while 60% struggle with inadequate computational resources, and 70% face difficulties managing large datasets. Despite these challenges, overcoming these issues will allow for the creation of "behaviomes", offering a deeper understanding of animal behavior. The Table 9 summarize the DL applications in animal behavior and pose estimation.

5. Challenges in implementing deep learning in animal farming

Despite its potential to revolutionize animal farming, DL faces several challenges that hinder widespread adoption. The complexity of deploying DL models in real-world farming environments arises from data limitations, environmental variability, computational constraints, and model generalization issues. Data scarcity and class imbalance make it difficult to train robust models, while inconsistent lighting, occlusion, and variations in animal appearance affect model accuracy. Additionally, high computational costs and the need for specialized hardware pose challenges for small and medium-scale farms. This section explores key obstacles in implementing DL for livestock monitoring, disease detection, behavior analysis, and farm automation, providing insights into potential solutions and future research directions.

5.1. Environmental variability

Animal farming environments are inherently dynamic, with lighting, weather, and spatial layouts creating variability that directly impacts the quality of data from sensors and monitoring systems. Deep learning (DL) models trained to identify behaviors may misinterpret low-light areas as abnormal activity or fail to detect subtle cues in overexposed zones. Outdoor farms face additional weather-related challenges, such as rain obscuring lenses, dust damaging sensors, or extreme temperatures disrupting data transmission (Neethirajan, 2017).

These factors collectively degrade the reliability of insights, risking delayed interventions for animal health or skewed sustainability metrics (Rencricea et al., 2023; Shulkin et al., 2025). Farm layouts further complicate monitoring, as localized microclimates require dense sensor networks to capture accurate readings. Animals in motion or partially obscured by objects pose additional challenges. Occlusion can hinder DL models in tracking individuals, affecting health monitoring, behavior analysis, and disease detection (Shulkin et al., 2025).

To address dynamic lighting inconsistencies, techniques such as data augmentation, high-quality imaging, infrared imaging, and adaptive contrast enhancement can improve DL tasks. Xu et al. (2020a) used Mask R-CNN to improve cattle detection accuracy under varying illumination by processing high-resolution drone imagery. Deep part-feature fusion (Hu et al., 2020) is combined with local and global visual cues to maintain identification robustness under occlusion and lighting changes. For behavior recognition, 3D CNNs and optical flow (Chen et al., 2021) is used to enhance motion detection sensitivity, capturing subtle livestock movements in low-light or overexposed conditions. In weather-affected environments, multi-sensor IoT systems (Niloofar et al., 2021) is integrated with thermal and visual data to compensate for rain or dust interference. In high-throughput phenotyping, Freitas et al. (2023) used a Mask R-CNN model to extract features despite environmental noise, ensuring reliable monitoring across diverse farming conditions.

5.2. Data scarcity and labeling issues

The scarcity of large, diverse datasets remains a critical bottleneck. Deep learning models require extensive labeled data, but collecting

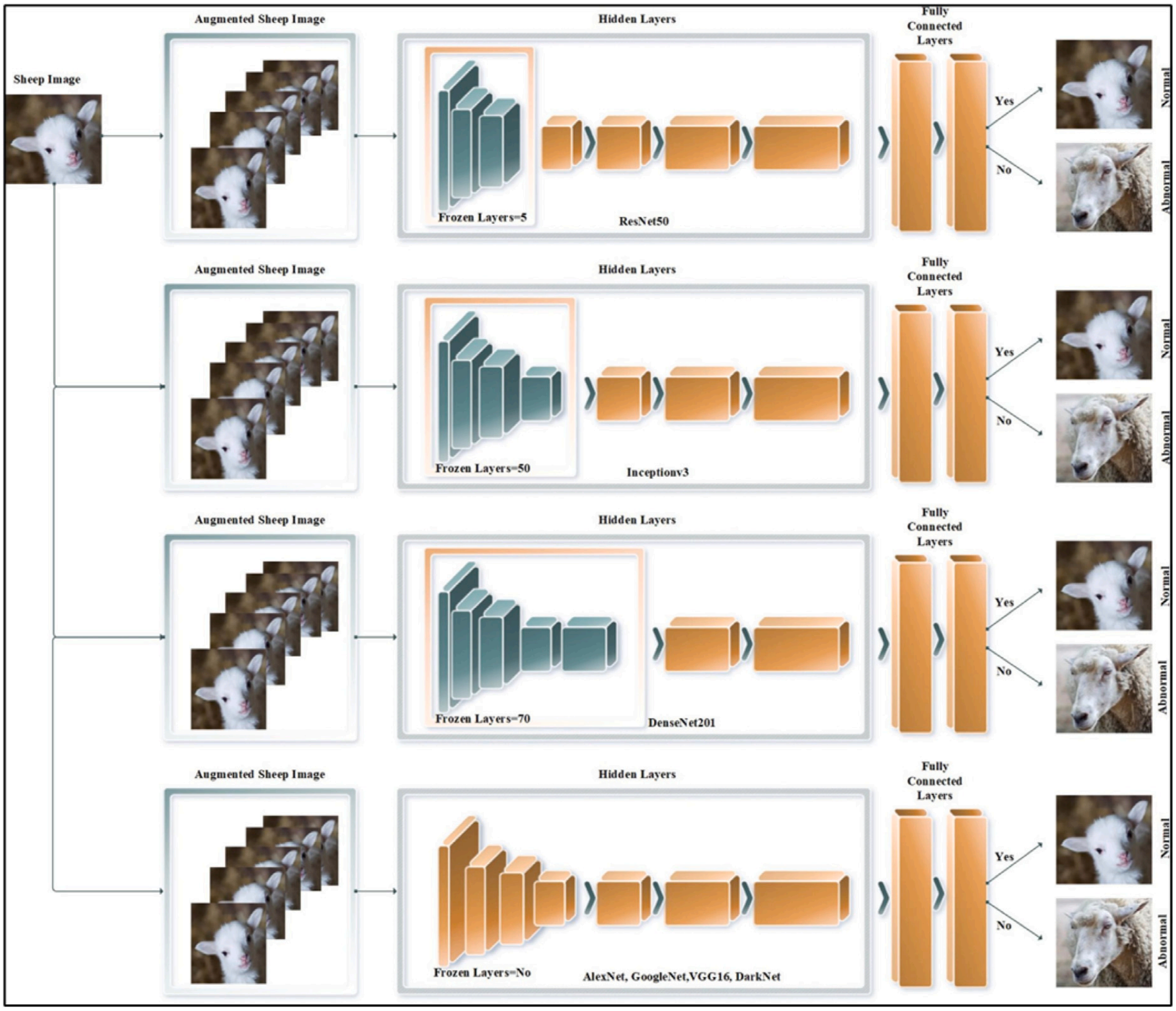

Fig. 25. Summary of all architecture's representation (Noor et al., 2020).

annotations for rare behaviors or species-specific conditions is labor-intensive and often requires veterinary expertise. For example, lighting conditions, occlusions from farm equipment, and differences in animal breeds introduce noise that complicates model training. Behaviors like lameness, grazing, or aggression requires specialized veterinary knowledge, which is costly and time-consuming to obtain (Pillay et al., 2024). Additionally, most existing datasets are small, proprietary, or restricted to controlled research settings, limiting their utility for real-world applications. Without standardized protocols for data collection and labeling, models trained on one dataset often fail to generalize across farms or even different herds within the same farm.

Annotation inconsistencies further exacerbate this issue, studies show inter-expert variability in labeling behaviors like grazing or aggression, reducing model reproducibility. To mitigate data scarcity, researchers are exploring synthetic data generation and collaborative learning approaches. Synthetic data augmentation, using techniques like 3D modeling or GANs, can expand small datasets by simulating livestock behaviors in virtual environments. However, these methods often fail to capture the complexity of real-world farm conditions, such as dirt, shadows, or interactions between animals. Federated learning, which allows multiple farms to train models collectively without sharing raw data, offers promise for privacy-preserving collaboration. Yet, variations in farm management practices, camera setups, and animal breeds can degrade model performance (Tjaudrasuwita et al., 2021). Self-supervised learning, which leverages unlabeled video to pre-train models, is another alternative, but it still requires fine-tuning with domain specific labels (Sharma et al., 2024a; Moosa et al., 2024; Wei et al., 2024). While these innovations help, their effectiveness depends on overcoming fundamental challenges: the need for larger, more diverse datasets and rigorous validation in real agricultural settings.

5.3. Real-time processing and scalability

Implementing deep learning for real-time animal monitoring such as health assessments or behavior tracking demands both high computational efficiency and low latency. Modern systems must process high resolution video streams from multiple cameras across large farms, requiring significant processing power to handle tasks like object detection or activity recognition (Li et al., 2022b). However, many farms lack the infrastructure for high performance computing, forcing reliance on edge devices with limited memory and energy constraints. Compounding this issue, algorithms must maintain accuracy despite environmental variables like changing lighting, occlusions, and unpredictable animal movements. For instance, while a high accuracy model might excel in post processing analysis, its computational complexity could introduce delays that render it impractical for real time alerts such as detecting early signs of illness that require immediate intervention. Striking this balance is critical, as even a few seconds of lag could mean the difference between preventing a disease outbreak or missing a critical warning sign.

To address these challenges, researchers are exploring lightweight neural architectures optimized for edge devices, which reduce computational load while preserving reasonable accuracy. Techniques like

Table 9
DL applications in animal behavior and pose estimation.

Author	Focus area	Methodology	Results	Applications
Yu et al. (2021)	Animal pose estimation	AP-10K dataset with images from 23 animal families	Advanced multi-species pose estimation; foundation for human pose model applications.	Multi-species pose estimation and training models for diverse species.
Lauer et al. (2022)	Enhancing DeepLabCut	Enhanced DeepLabCut for occlusion handling	Improved tracking accuracy; addressed occlusion and appearance similarity.	Animal behavior studies, particularly in complex settings.
Wang et al. (2021)	Livestock weight prediction	Semantic segmentation, regression models like ResNet-101-D	Errors as low as 13.11 pounds in weight prediction.	Non-invasive livestock weight prediction in farming.
Li and Lee (2021)	Animal pose estimation	UDA approach with MDAM and student–teacher network	Improved accuracy and generalization across species using TigDog and VisDA 2019 datasets.	Animal pose estimation with better generalization for diverse species.
Chen et al. (2021)	Livestock behavior recognition	CNNs and RNNs for behavior recognition in livestock	Superior performance in diverse environments for real-time livestock welfare and health monitoring.	Livestock welfare and health monitoring in farming environments.
Pereira et al. (2022)	Multi-animal pose tracking	SLEAP multi-animal pose tracking framework with machine learning techniques	Achieved <2.8 pixel error on 95% of points and processed at 320 FPS, revolutionizing ecological monitoring.	Advanced animal behavior analysis and ecological monitoring.
Achour et al. (2020)	Dairy cow identification and behavior	CNN-based system for dairy cow behavior and feeding monitoring	Achieved 97.95% accuracy, with fast training and low computational cost.	Real-time dairy cow management and feeding behavior monitoring.
Fuentes et al. (2020)	Cattle behavior recognition	CNNs with spatio-temporal information and data augmentation	Achieved over 95% precision, enhancing real-time health monitoring for cattle welfare.	Real-time cattle health and welfare monitoring for farming.
Manoharan (2020)	Dairy cow behavior recognition	Res-DenseYOLO for behavior recognition, particularly feeding behavior	Achieved 94.7% precision and 96.3% mAP, with over 98% accuracy in feeding behavior detection.	Health monitoring and feeding behavior detection in dairy cows.
Jiang et al. (2020)	Goat behavior recognition	IMUs for movement tracking, YOLOv5 for counting and tracking	Achieved 92.19% mAP, providing real-time insights into goat health and behavior.	Goat health and behavior monitoring for farming and welfare management.
Noor et al. (2020)	Sheep facial expression classification	CNNs for facial expression classification (stress, happiness)	Achieved >90% accuracy in detecting emotional states, improving welfare management.	Sheep welfare monitoring through facial expression analysis.
Fang et al. (2021)	Broiler chicken behavior monitoring	CNNs and data augmentation for pose estimation and behavior monitoring	Improved real-time performance and pose estimation for poultry welfare management.	Poultry welfare management through behavior monitoring and real-time insights.
Nasiri et al. (2024b)	Broiler behavior quantification	Automated system using behavior recognition for welfare monitoring	Achieved 96.7% accuracy and 92.4% AUC for monitoring behaviors like stretching and preening.	Broiler welfare behavior monitoring for improved farming practices.
Rickert et al. (2020)	Pig posture detection	Lightweight CNN with attention mechanisms for posture detection	Achieved 97.7% accuracy for standing pigs and 84.1% for mounting pigs, with 63 ms inference speed.	Continuous monitoring of pig behavior and welfare in farms.

model quantization and pruning further streamline inference speed without drastic performance drops (Chelotti et al., 2023). Another approach involves hierarchical processing, where simple algorithms filter out irrelevant frames before applying more complex models to high-priority scenes. However, such optimizations risk overlooking subtle but critical behaviors for example, a lightweight model might miss early signs of respiratory distress if trained only on coarse features. Cloud edge hybrid systems offer a compromise, offloading intensive tasks to centralized servers while edge devices handle time-sensitive decisions. Yet, this depends on reliable connectivity, which is often unavailable in rural farming environments. Ultimately, achieving real time efficiency without sacrificing accuracy remains an open problem, necessitating domain specific adaptations tailored to the constraints of agricultural settings.

5.4. Generalization across different species

Deep learning models often struggle when applied to new farm environments or animal species. Trained on specific conditions, they fail with pasture systems or different species due to lighting, layout, and biological variations. A Holstein-focused lameness detector may misclassify Jerseys, while pig behavior models underperform in free range settings. This lack of adaptability forces expensive retraining for each scenario. Even minor changes in camera angles or animal density can degrade performance, limiting real world usability (Fuentes et al., 2022).

Domain adaptation and synthetic data help bridge generalization gaps. Multi species training teaches models universal animal features, while simulated datasets expand environmental coverage. However, synthetic data often misses real-world nuances like weather effects. Federated learning enables cross farm collaboration without sharing raw data, but implementation remains challenging (Kappes et al., 2023). Future solutions may combine these approaches with standardized testing protocols to ensure reliability across diverse farming conditions.

6. Future directions and recommendations

The implementation of deep learning in animal farming holds significant promise, but several challenges must be addressed to maximize its potential. A key issue is the scarcity of high-quality, annotated datasets, which can be mitigated by improving data collection methods, such as the use of automated sensors, cameras, and wearable devices (Selle et al., 2023) to solve this issue deploy automated sensors, cameras, and wearable devices to gather continuous, high-resolution behavioral and physiological data. Use multi-modal sensing (Infrared, thermal, RFID) to cross-validate annotations and reduce labeling errors. Additionally, collaboration between farmers and AI developers is essential to create standardized protocols for data annotation.

To make deep learning more accessible and scalable, cost-effective solutions are needed especially for smaller farms. This can be achieved through the development of affordable sensors and open-source AI frameworks. Collaboration with regulatory bodies is also essential to establish clear guidelines for AI deployment, ensuring safety and compliance with animal welfare standards (Fazzari et al., 2024). The deep learning models must be capable of continuous learning and updating to adapt to changing farming conditions, emerging diseases, and new technologies.

Deep learning can revolutionize the way farmers track livestock behavior, detect diseases early, and optimize breeding programs through precise, data-driven insights (Korkmaz et al., 2024). By automating routine tasks such as monitoring animal welfare or identifying illnesses, AI has the potential to reduce human error, improve efficiency, and enhance animal well-being. It is essential to develop more robust, adaptable models capable of generalizing across diverse farming tasks (segmentation, classification, and localization).

Future advancements should focus on greater integration of AI with existing farm management systems to streamline operations and improve animal welfare. This includes developing user-friendly interfaces and decision support tools that allow farm managers to easily interact with AI systems (Bhujel et al., 2024). Transparent decision-making processes and the involvement of veterinary experts in system development will ensure that AI-based interventions are aligned with ethical standards. Furthermore, handling issues like occlusion in visual recognition systems can be improved through advanced imaging techniques and multi-camera setups. Moreover, addressing real-time processing challenges through edge computing and optimized algorithms is crucial for applications such as health monitoring and behavior tracking.

For long term and short term advancements plaining in deep learning for animal farming are expected to unfold across short, medium, and long term horizons. In the short term (1–3 years), we anticipate the proliferation of edge computing devices tailored for farm environments, enabling real-time animal monitoring even in areas with limited internet access. These systems will increasingly incorporate multimodal sensing for holistic animal behavior analysis (Jobarteh et al., 2024; Chae et al., 2024). Concurrently, transfer learning techniques will become more robust, offering pre-trained models optimized for agricultural data scarcity and specific livestock contexts. High quality synthetic data generation, such as through GANs and generative algorithms, will also see growth as a remedy to the challenges posed by limited high quality datasets. Federated learning models may emerge as a privacy conscious solution, allowing multiple farms to contribute to a shared model without relinquishing sensitive data (Arshad et al., 2024; Ghasemkhani et al., 2024). In traditional deep learning, data from various sources must be centralized in a single server or cloud environment, which raises serious concerns about data privacy, ownership, and security especially for smallholder farmers and private agricultural enterprises. Federated learning overcomes this by allowing machine learning models to be trained across multiple decentralized devices or farms while keeping the data localized.

In the medium to long term (3–5 years and beyond), deep learning is likely to be integrated with complementary technologies such as blockchain for traceability, augmented reality for guided animal care, and autonomous robotics for farm automation. Cross-species learning and generalizable AI models could enable knowledge transfer between livestock types, improving adaptability across farming systems. As regulatory frameworks become more defined, explainable AI will gain traction, ensuring that decisions in livestock health and welfare can be interpreted and justified (Das et al., 2024; Aqib et al., 2023; Yogi et al., 2024). The fusion of phenotypic, genomic, and behavioral data will revolutionize precision breeding and welfare assessments, responding to growing consumer and ethical concerns. Finally, scalable solutions will be developed to enable adoption by small and medium-sized farms, supported by training tools and low-cost AI platforms, fostering more equitable access to smart agriculture advancements.

7. Conclusion

In conclusion, deep learning holds significant promise for revolutionizing animal farming by improving productivity, precision in animal health monitoring, and decision-making processes. However, the successful implementation of AI in this field faces several challenges, such as data scarcity, environmental variability, and the need for real-time processing capabilities. Addressing these challenges will require enhanced data collection strategies, the development of more efficient algorithms, and the integration of domain-specific expertise, such as veterinary knowledge, to ensure that AI systems align with animal welfare standards. Additionally, overcoming issues like occlusion, noise, and generalization across diverse environments will be crucial for building robust AI models that can be applied across different farming contexts. By fostering collaboration between researchers, farmers, and technology developers, as well as investing in scalable and ethical solutions, deep learning can offer transformative benefits for the future of animal farming. With continued innovation and attention to the practical aspects of deployment, AI has the potential to significantly improve both the economic and ethical dimensions of animal agriculture.

Advancements in computer vision and predictive analytics are reshaping livestock management. While RGB-D sensors enable 3D biometric measurements, their dependency on stable lighting and high data storage costs remain challenges. Blockchain technology enhances data traceability, ensuring accountability in environmental monitoring and regulatory compliance. Mobile applications and cloud-based platforms streamline real-time access to sensor data, facilitating quick interventions for issues like heat stress or disease outbreaks. However, scalability, interoperability, and ethical concerns (e.g., privacy in camera systems) persist, necessitating standardized frameworks for data governance and security. By addressing these gaps, precision livestock farming can balance animal welfare, sustainability, and economic viability in increasingly complex environmental conditions.

To address these challenges, researchers are adopting multi-modal sensor fusion and transfer learning. IoT systems integrating thermal cameras, accelerometers, and gas sensors enable cross-validation of environmental and biometric data, improving robustness against single source failures (Dufourq et al., 2022; Berckmans, 2017). Transfer learning leverages pre-trained models from human medical imaging or crop monitoring, adapting them to livestock contexts with limited data, though this requires careful fine-tuning to account for species specific anatomies. Advances in automated annotation tools, such as AI-assisted bounding boxes for tracking animal movements, reduce labeling costs by up to 60% while maintaining accuracy. Future efforts must prioritize standardized benchmarking datasets and regulatory frameworks to ensure AI models meet clinical utility standards in diverse farming conditions.

CRediT authorship contribution statement

Zahid Ur Rahman: Writing – original draft, Visualization, Methodology, Investigation, Formal analysis. **Mohd Shahrimie Mohd Asaari:** Writing – review & editing, Validation, Supervision, Resources, Project administration, Investigation, Funding acquisition, Conceptualization. **Haidi Ibrahim:** Writing – review & editing, Validation, Supervision.

Declaration of competing interest

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

Acknowledgment

This research is supported by Ministry of Higher Education Malaysia (MOHE) Fundamental Research Grant Scheme FRGS/1/2023/TK07/USM/02/6.

Data availability

No data was used for the research described in the article.

References

Abangan, A.S., Kopp, D., Faillettaz, R., 2023. Artificial intelligence for fish behavior recognition may unlock fishing gear selectivity. Front. Mar. Sci. 10.

Achour, B., Belkadi, M., Filali, I., Laghrouche, M., Lahdir, M., 2020. Image analysis for individual identification and feeding behaviour monitoring of dairy cows based on convolutional neural networks (cnn). Biosyst. Eng. 198, 31–49.

Aguilar-Lazcano, C.A., Espinosa-Curiel, I.E., Ríos-Martínez, J.A., Madera-Ramírez, F.A., Pérez-Espinosa, H., 2023. Machine learning-based sensor data fusion for animal monitoring: Scoping review. Sensors 23 (12), 5732.

Ahmod, G., Malick, R.A.S., Akhunzada, A., Zahid, S., Sagri, M.R., Gani, A., 2021. An approach towards iot-based predictive service for early detection of diseases in poultry chickens. Sustain. (Switzerland) 13 (23).

Akcay, H.G., Kabasakal, B., Aksu, D., Demir, N., Oz, M., Erdogan, A., 2020. Automated bird counting with deep learning for regional bird distribution mapping. Animals 10 (7), 1207.

Akkem, Y., Biswas, S.K., Varanasi, A., 2024. A comprehensive review of synthetic data generation in smart farming by using variational autoencoder and generative adversarial network. Eng. Appl. Artif. Intell. 131, 107881.

Alameer, A., Kyriazakis, I., Bacardit, J., 2020. Automated recognition of postures and drinking behaviour for the detection of compromised health in pigs. Sci. Rep. 10 (1).

Andrew, W., Gao, J., Mullan, S., Campbell, N., Dowsey, A.W., Burghardt, T., 2021. Visual identification of individual holstein-friesian cattle via deep metric learning. Comput. Electron. Agric. 185, 106133.

Aqib, A.I., Fatima, M., Muneer, A., Atta, K., Arslan, M., Zaheer, C.-N.F., Muneer, S., Murtaza, M., 2023. Explainable artificial intelligence (XAI) in the veterinary and animal sciences field. In: Explainable Artificial Intelligence for Biomedical Applications. River Publishers, pp. 33–56.

Arshad, J., Irtisam, A., Arif, T., Rasheed, M.S., Chauhdary, S.T., Rahmani, M.K.I., Almajalid, R., 2024. A federated learning model for intelligent cattle health monitoring system using body area sensors and IoT. Egypt. Inform. J. 27, 100488.

Bao, K., Bi, J., Ma, R., Sun, Y., Zhang, W., Wang, Y., 2023. A spatial-reduction attention-based bigru network for water level prediction. Water 15 (7), 1306.

Berckmans, D., 2017. General introduction to precision livestock farming. Anim. Front. 7 (1), 6–11.

Besson, M., Alison, J., Bjerge, K., Gorochowski, T.E., Høye, T.T., Jucker, T., Mann, H.M.R., Clements, C.F., 2022. Towards the fully automated monitoring of ecological communities. Ecol. Lett. 25 (12), 2753–2775.

Bezen, R., Edan, Y., Halachmi, I., 2020. Computer vision system for measuring individual cow feed intake using RGB-D camera and deep learning algorithms. Comput. Electron. Agric. 172, 105345.

Bhujel, A., Wang, Y., Lu, Y., Morris, D., Dangol, M., 2024. Public computer vision datasets for precision livestock farming: A systematic survey. arXiv preprint arXiv:2406.10628.

Biswas, A., 2024. Media insights engine for advanced media analysis: A case study of a computer vision innovation for pet health diagnosis. arXiv preprint arXiv:2407.13679.

Bochkovskiy, A., Wang, C.Y., Liao, H.Y.M., 2020. Yolov4: Optimal speed and accuracy of object detection. arXiv preprint arXiv:2004.10934.

Bondi, E., Jain, R., Aggrawal, P., Anand, S., Hannaford, R., Kapoor, A., Piavis, J., Shah, S., Joppa, L., Dilkina, B., et al., 2020. Birdsai: A dataset for detection and tracking in aerial thermal infrared videos. In: Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision. pp. 1747–1756.

Campos, R.L., Yoon, S.C., Chung, S., Bhandarkar, S.M., 2023. Semi-supervised deep learning for the detection of foreign materials on poultry meat with near-infrared hyperspectral imaging. Sensors 23 (16).

Chae, H., Lee, J., Kim, J., Lee, S., Lee, J., Chung, Y., Park, D., 2024. Novel method for detecting coughing pigs with audio-visual multimodality for smart agriculture monitoring. Sensors 24 (7232).

Chelotti, J., Martinez-Rau, L., Ferrero, M., Vignolo, L., Galli, J., Planisich, A., Rufiner, H.L., Giovanini, L., 2023. Livestock feeding behavior: A tutorial review on automated techniques for ruminant monitoring. arXiv preprint arXiv:2312.09259.

Chen, B.-L., Cheng, T.-H., Huang, Y.-C., Hsieh, Y.-L., Hsu, H.-C., Lu, C.-Y., Huang, M.-H., Nien, S.-Y., Kuo, Y.-F., 2023a. Developing an automatic warning system for anomalous chicken dispersion and movement using deep learning and machine learning. Poult. Sci. 102 (12), 103040.

Chen, Z., Lu, J., Wang, H., 2023c. A review of posture detection methods for pigs using deep learning. Appl. Sci. 13 (12), 6997.

Chen, X., Roberts, R., Liu, Z., Tong, W., 2023b. A generative adversarial network model alternative to animal studies for clinical pathology assessment. Nat. Commun. 14 (1).

Chen, C., Zhu, W., Norton, T., 2021. Behaviour recognition of pigs and cattle: Journey from computer vision to deep learning. Comput. Electron. Agric. 187, 106255.

Chen, C., Zhu, W., Steibel, J., Siegford, J., Han, J., Norton, T., 2020a. Classification of drinking and drinker-playing in pigs by a video-based deep learning method. Biosyst. Eng. 196, 1–14.

Chen, C., Zhu, W., Steibel, J., Siegford, J., Han, J., Norton, T., 2020b. Recognition of feeding behaviour of pigs and determination of feeding time of each pig by a video-based deep learning method. Comput. Electron. Agric. 176, 105642.

Chollet, F., 2017. Xception: Deep learning with depthwise separable convolutions. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. pp. 1251–1258.

Clapham, M., Miller, E., Nguyen, M., Darimont, C.T., 2020. Automated facial recognition for wildlife that lack unique markings: A deep learning approach for brown bears. Ecol. Evol. 10 (23), 12883–12892.

Corcoran, E., Winsen, M., Sudholz, A., Hamilton, G., 2021. Automated detection of wildlife using drones: Synthesis, opportunities and constraints. Methods Ecol. Evol. 12 (6), 1103–1114.

Cui, J., Li, K., Lv, Y., Liu, S., Cai, Z., Luo, R., Zhang, Z., Wang, S., 2024. Development of a new hyperspectral imaging technology with autoencoder-assisted generative adversarial network for predicting the content of polyunsaturated fatty acids in red meat. Comput. Electron. Agric. 220, 108842.

Darma, I.N.Y.P., Siagian, S.N.F., Darwin, R.A., Sihotang, E.F.A., Irwansyah, E., 2023. Implementation of convolutional neural network to minimize wildlife-vehicle collisions. In: 2023 5th International Conference on Cybernetics and Intelligent System. ICORIS, IEEE, pp. 1–5.

Das, N., Padhy, N., Dey, N., Paul, H., Chowdhury, S., 2024. Exploring explainable AI methods for bird sound-based species recognition systems. Multimedia Tools Appl. 83 (24), 64223–64253.

Degu, M.Z., Simegn, G.L., 2023. Smartphone based detection and classification of poultry diseases from chicken fecal images using deep learning techniques. Smart Agric. Technol. 4, 100221.

Ditria, E.M., Lopez-Marcano, S., Sievers, M., Jinks, E.L., Brown, C.J., Connolly, R.M., 2020. Automating the analysis of fish abundance using object detection: Optimizing animal ecology with deep learning. Front. Mar. Sci. 7.

Duan, E., Han, G., Zhao, S., Ma, Y., Lv, Y., Bai, Z., 2023. Regulation of meat duck activeness through photoperiod based on deep learning. Animals 13 (22), 3520.

Dufourq, E., Batist, C., Foquet, R., Durbach, I., 2022. Passive acoustic monitoring of animal populations with transfer learning. Ecol. Inform. 70, 101688.

Duporge, I., Isupova, O., Reece, S., Macdonald, D.W., Wang, T., 2021. Using very-high-resolution satellite imagery and deep learning to detect and count African elephants in heterogeneous landscapes. Remote. Sens. Ecol. Conserv. 7 (3), 369–381.

Elmessery, W.M., Gutiérrez, J., Abd El-Wahhab, G.G., Elkhaiat, I.A., El-Soaly, I.S., Alhag, S.K., Al-Shuraym, L.A., Akela, M.A., Moghanm, F.S., Abdelshafie, M.F., 2023. Yolo-based model for automatic detection of broiler pathological phenomena through visual and thermal images in intensive poultry houses. Agriculture 13 (8), 1527.

Fan, D.-P., Ji, G.-P., Sun, G., Cheng, M.-M., Shen, J., Shao, L., 2020. Camouflaged object detection. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 2777–2787.

Fang, C., Zhang, T., Zheng, H., Huang, J., Cuan, K., 2021. Pose estimation and behavior classification of broiler chickens based on deep neural networks. Comput. Electron. Agric. 180, 105863.

Fazzari, E., F, D.R., Stefanini, C., 2024. Animal behavior analysis methods using deep learning: A survey. arXiv preprint arXiv:2405.14002.

Feighelstein, M., Ehrlich, Y., Naftaly, L., Alpin, M., Nadir, S., Shimshoni, I., Pinho, R.H., Luna, S.P.L., Zamansky, A., 2023. Deep learning for video-based automated pain recognition in rabbits. Sci. Rep. 13 (1).

Ferreira, A.C., Silva, L.R., Renna, F., Brandl, H.B., Renoult, J.P., Farine, D.R., Covas, R., Doutrelant, C., 2020. Deep learning-based methods for individual recognition in small birds. Methods Ecol. Evol. 11 (9), 1072–1085.

Freitas, M.V., Lemos, C.G., Ariede, R.B., Agudelo, J.F., Neto, R.R., Borges, C.H., Mastrochirico-Filho, V.A., Porto-Foresti, F., Iope, R.L., Batista, F.M., et al., 2023. High-throughput phenotyping by deep learning to include body shape in the breeding program of pacu (piaractus mesopotamicus). Aquaculture 562, 738847.

Fuentes, S., Gonzalez Viejo, C., Tongson, E., Dunshea, F.R., 2022. The livestock farming digital transformation: Implementation of new and emerging technologies using artificial intelligence. Anim. Heal. Res. Rev. 23 (1), 59–71.

Fuentes, A., Yoon, S., Park, J., Park, D.S., 2020. Deep learning-based hierarchical cattle behavior recognition with spatio-temporal information. Comput. Electron. Agric. 177, 105627.
Gao, T., Li, J., Watanabe, Y., Hung, C., Yamanaka, A., Horie, K., Yanagisawa, M., Ohsawa, M., Kume, K., 2021. GI-SleepNet: A highly versatile image-based sleep classification using a deep learning algorithm. Clocks Sleep 3 (4), 581–597.
García, R., Aguilar, J., Toro, M., Pinto, A., Rodríguez, P., 2020. A systematic literature review on the use of machine learning in precision livestock farming. Comput. Electron. Agric. 179, 105826.
Ghasemkhani, B., Varliklar, O., Dogan, Y., Utku, S., Birant, K.U., Birant, D., 2024. Federated multi-label learning (FMLL): Innovative method for classification tasks in animal science. Animals 14 (14), 2021.
Goodfellow, I.J., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., Bengio, Y., 2014. Generative adversarial nets. Adv. Neural Inf. Process. Syst. 27.
Goodwin, M., Halvorsen, K.T., Jiao, L., Knausgård, K.M., Martin, A.H., Moyano, M., Oomen, R.A., Rasmussen, J.H., Sørdalen, T.K., Thorbjørnsen, S.H., 2022. Unlocking the potential of deep learning for marine ecology: Overview, applications, and outlook. ICES J. Mar. Sci. 79 (2), 319–336.
Guarnido-Lopez, P., Ramirez-Agudelo, J.-F., Denimal, E., Benaouda, M., 2024. Programming and setting up the object detection algorithm YOLO to determine feeding activities of beef cattle: A comparison between YOLOv8m and YOLOv10m. Animals 14 (19), 2821.
Guerri, M.F., Distante, C., Spagnolo, P., Bougouri, F., Taleb-Ahmed, A., 2024. Deep learning techniques for hyperspectral image analysis in agriculture: A review. ISPRS Open J. Photogramm. Remote. Sens. 100062.
Gunda, V.S.P., Gulla, H., Kosana, V., Janapati, S., 2022. A hybrid deep learning based robust framework for cattle identification. In: ASSIC 2022 - Proceedings: International Conference on Advancements in Smart, Secure and Intelligent Computing. Bhubaneswar, India, pp. 1–5.
Gupta, G., Kshirsagar, M., Zhong, M., Gholami, S., Ferres, J.L., 2021. Comparing recurrent convolutional neural networks for large scale bird species classification. Sci. Rep. 11 (1).
Han, F., Yao, J., Zhu, H., Wang, C., 2020. Marine organism detection and classification from underwater vision based on the deep CNN method. Math. Probl. Eng. 2020, 1–11.
Hazra, D., Kim, M.R., Byun, Y.C., 2022. Generative adversarial networks for creating synthetic nucleic acid sequences of cat genome. Int. J. Mol. Sci. 23 (7), 3701.
He, K., Zhang, X., Ren, S., Sun, J., 2015. Deep residual learning for image recognition. arXiv preprint arXiv:1512.03385.
Hossain, M.E., Kabir, M.A., Zheng, L., Swain, D.L., McGrath, S., Medway, J., 2022. A systematic review of machine learning techniques for cattle identification: Datasets, methods and future directions. Artif. Intell. Agric. 6, 138–155.
Hou, J., He, Y., Yang, H., Connor, T., Gao, J., Wang, Y., Zeng, Y., Zhang, J., Huang, J., Zheng, B., et al., 2020. Identification of animal individuals using deep learning: A case study of giant panda. Biol. Cons. 242, 108414.
Howard, A.G., Zhu, M., Chen, B., Kalenichenko, D., Wang, W., Weyand, T., Andreetto, M., Adam, H., 2017. Mobilenets: Efficient convolutional neural networks for mobile vision applications. arXiv preprint arXiv:1704.04861.
Hu, H., Dai, B., Shen, W., Wei, X., Sun, J., Li, R., Zhang, Y., 2020. Cow identification based on fusion of deep parts features. Biosyst. Eng. 192, 245–256.
Huang, E., He, Z., Mao, A., Camila Ceballos, M., Parsons, T.D., Liu, K., 2023. A semi-supervised generative adversarial network for amodal instance segmentation of piglets in farrowing pens. Comput. Electron. Agric. 209.
Ibraheam, M., Gebali, F., Li, K.F., Sielecki, L., 2020. Animal species recognition using deep learning. In: Advanced Information Networking and Applications. Vol. 1151, pp. 629–640.
Islam, S.B., Valles, D., Hibbitts, T.J., Ryberg, W.A., Walkup, D.K., Forstner, M.R.J., 2023. Animal species recognition with deep convolutional neural networks from ecological camera trap images. Animals 13, 1526.
Jalal, A., Salman, A., Mian, A., Shortis, M., Shafait, F., 2020. Fish detection and species classification in underwater environments using deep learning with temporal information. Ecol. Inform. 57, 101088.
Jegham, N., Koh, C.Y., Abdelatti, M., Hendawi, A., 2024. Evaluating the evolution of yolo (you only look once) models: A comprehensive benchmark study of yolo11 and its predecessors. arXiv preprint arXiv:2411.00201.
Jiang, M., Rao, Y., Zhang, J., Shen, Y., 2020. Automatic behavior recognition of group-housed goats using deep learning. Comput. Electron. Agric. 177, 105706.
Jobarteh, B., Mincu, M., Dinu, G., Neethirajan, S., 2024. Multi modal information fusion of acoustic and linguistic data for decoding dairy cow vocalizations in animal welfare assessment. arXiv preprint arXiv:2411.00477.
Jocher, G., Stoken, A., Borovec, J., et al., 2020. Ultralytics/yolov5: v3.1 - bug fixes and performance improvements.
Kahl, S., Wood, C.M., Eibl, M., Klinck, H., 2021. Birdnet: A deep learning solution for avian diversity monitoring. Ecol. Inform. 61, 101236.
Kandimalla, V., Richard, M., Smith, F., Quirion, J., Torgo, L., Whidden, C., 2022. Automated detection, classification and counting of fish in fish passages with deep learning. Front. Mar. Sci. 8.
Kappes, A., Tozooneyi, T., Shakil, G., Railey, A.F., McIntyre, K.M., Mayberry, D.E., Rushton, J., Pendell, D.L., Marsh, T.L., 2023. Livestock health and disease economics: A scoping review of selected literature. Front. Vet. Sci. 10.
Kathir, M., Balaji, V., Ashwini, K., 2024. Animal intrusion detection using yolo V8. In: 2024 10th International Conference on Advanced Computing and Communication Systems. ICACCS, pp. 206–211.
Katija, K., Roberts, P.L.D., Daniels, J., Lapides, A., Barnard, K., Risi, M., Ranaan, B.Y., Woodward, B.G., Takahashi, J., 2021. Visual tracking of deepwater animals using machine learning-controlled robotic underwater vehicles. In: 2021 IEEE Winter Conference on Applications of Computer Vision. WACV, pp. 859–868.
Khanal, R., Choi, Y., Lee, J., 2024. Transforming poultry farming: A pyramid vision transformer approach for accurate chicken counting in smart farm environments. Sensors 24 (10), 2977.
Khanam, R., Hussain, M., 2024. Yolov11: An overview of the key architectural enhancements. arXiv preprint arXiv:2410.17725.
Khatri, K., Asha, C.C., D'Souza, J.M., 2022. Detection of animals in thermal imagery for surveillance using GAN and object detection framework. In: 2022 International Conference for Advancement in Technology. ICONAT 2022, IEEE, Goa, India, pp. 1–6.
Kim, E., Moon, J., Shim, J., Hwang, E., 2023. DualDiscWaveGAN-based data augmentation scheme for animal sound classification. Sensors 23 (4), 2024.
Kingma, D.P., Welling, M., et al., 2013. Auto-encoding variational bayes.
Kingma, D.P., Welling, M., et al., 2019. An introduction to variational autoencoders. Found. Trends Mach. Learn. 12 (4), 307–392.
Knausgård, K.M., Wiklund, A., Sørdalen, T.K., Halvorsen, K.T., Kleiven, A.R., Jiao, L., Goodwin, M., 2022. Temperate fish detection and classification: A deep learning based approach. Appl. Intell. 52 (6), 6988–7001.
Kopets, E., Shpilevaya, T., Vasilchenko, O., Karimov, A., Butusov, D., 2024. Generating synthetic sperm whale voice data using StyleGAN2-ADA. Big Data Cogn. Comput. 8 (4), 40.
Korkmaz, A., Agdas, M.T., Kosunalp, S., Iliev, T., Stoyanov, I., 2024. Detection of threats to farm animals using deep learning models: A comparative study. Appl. Sci. 14 (14), 6098.
Krizhevsky, A., Sutskever, I., Hinton, G.E., 2012. Imagenet classification with deep convolutional neural networks. In: Advances in Neural Information Processing Systems. Vol. 25.
Lauer, J., Zhou, M., Ye, S., Menegas, W., Schneider, S., Nath, T., Rahman, M.M., Di Santo, V., Soberanes, D., Feng, G., et al., 2022. Multi-animal pose estimation, identification and tracking with DeepLabCut. Nature Methods 19 (4), 496–504.
LeCun, Y., Bottou, L., Bengio, Y., Haffner, P., 1998. Gradient-based learning applied to document recognition. Proc. IEEE 86 (11), 2278–2324.
Li, J., Green-Miller, A.R., Hu, X., Lucic, A., Mohan, M.M., Dilger, R.N., Condotta, I.C., Aldridge, B., Hart, J.M., Ahuja, N., 2022b. Barriers to computer vision applications in pig production facilities. Comput. Electron. Agric. 200, 107227.
Li, C., Lee, G.H., 2021. From synthetic to real: Unsupervised domain adaptation for animal pose estimation. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition.
Li, C., Li, L., Jiang, H., Weng, K., Geng, Y., Li, L., Ke, Z., Li, Q., Cheng, M., Nie, W., et al., 2022a. Yolov6: A single-stage object detection framework for industrial applications. arXiv preprint arXiv:2209.02976.
Li, K., Teng, G., Wang, J., Zhang, Y., Gao, L., Feng, H., 2024. Body condition scoring of dairy cows based on feature point location. IEEE Access 12, 5270–5283.
Li, J., Xu, W., Deng, L., Xiao, Y., Han, Z., Zheng, H., 2023. Deep learning for visual recognition and detection of aquatic animals: A review. Rev. Aquac. 15 (2), 409–433.
Liakos, K.G., Busato, P., Moshou, D., Pearson, S., Bochtis, D., 2018. Machine learning in agriculture: A review. Sensors 18 (8).
Liu, Y., Li, W., Liu, X., Li, Z., Yue, J., 2024. Deep learning in multiple animal tracking: A survey. Comput. Electron. Agric. 224, 109161.
Liu, D., Oczak, M., Maschat, K., Baumgartner, J., Pletzer, B., He, D., Norton, T., 2020. A computer vision-based method for spatial–temporal action recognition of tail-biting behaviour in group-housed pigs. Biosyst. Eng. 195, 27–41.
Lu, Z., Zhao, M., Luo, J., Wang, G., Wang, D., 2021. Automatic teat detection for rotary milking system based on deep learning algorithms. Comput. Electron. Agric. 189, 106391.
Luo, S., Ma, Y., Jiang, F., Wang, H., Tong, Q., Wang, L., 2023. Dead laying hens detection using TIR-nir-depth images and deep learning on a commercial farm. Animals 13 (11).
Lv, Y., Zhang, J., Dai, Y., Li, A., Liu, B., Barnes, N., Fan, D.-P., 2021. Simultaneously localize, segment and rank the camouflaged objects. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 11591–11601.
Ma, C., Deng, M., Yin, Y., 2024. Pig face recognition based on improved YOLOv4 lightweight neural network. Inf. Process. Agric. 11 (3), 356–371.
Ma, X., Lu, X., Huang, Y., Yang, X., Xu, Z., Mo, G., Ren, Y., Li, L., 2022. An advanced chicken face detection network based on GAN and MAE. Animals 12 (21), 3055.
Madhumathi, C.S., Naveen, V., Akshay, N., Kumar, M.S., Aslam, M.M., 2023. Advanced wild animal detection and alert system using YOLO V5 model. In: 2023 7th International Conference on Trends in Electronics and Informatics. ICOEI, pp. 365–371.
Manoharan, D.S., 2020. Embedded imaging system based behavior analysis of dairy cow.

Mao, A., Giraudet, C.S.E., Liu, K., De Almeida Nolasco, I., Xie, Z., Xie, Z., Gao, Y., Theobald, J., Bhatta, D., Stewart, R., et al., 2022. Automated identification of chicken distress vocalizations using deep learning models. J. R. Soc. Interface 19 (191).

Martin-Abadal, M., Ruiz-Frau, A., Hinz, H., Gonzalez-Cid, Y., 2020. Jellytoring: Real-time jellyfish monitoring based on deep learning object detection. Sensors 20 (6), 1708.

Meckbach, C., Tiesmeyer, V., Traulsen, I., 2021. A promising approach towards precise animal weight monitoring using convolutional neural networks. Comput. Electron. Agric. 183, 106056.

Mirza, M., Osindero, S., 2014. Conditional generative adversarial nets. arXiv preprint arXiv:1411.1784.

Mitra, J., Qiu, J., MacDonald, M., Venugopal, P., Wallace, K., Abdou, H., Richmond, M., Elansary, N., Edwards, J., Patel, N., et al., 2022. Automatic hemorrhage detection from color doppler ultrasound using a generative adversarial network (GAN)-based anomaly detection method. IEEE J. Transl. Eng. Heal. Med. 10, 1–9.

Moosa, M., Cheikh, F.A., Beghdadi, A., Ullah, M., 2024. Self-supervised animal detection, tracking & re-identification. In: 2024 IEEE Thirteenth International Conference on Image Processing Theory, Tools and Applications. IPTA, pp. 01–06.

Nakrosis, A., Paulauskaite-Taraseviciene, A., Raudonis, V., Narusis, I., Gruzauskas, V., Gruzauskas, R., Lagzdinyte-Budnike, I., 2023. Towards early poultry health prediction through non-invasive and computer vision-based dropping classification. Animals 13 (19), 3041.

Nasiri, A., Amirivojdan, A., Zhao, Y., Gan, H., 2024a. An automated video action recognition-based system for drinking time estimation of individual broilers. Smart Agric. Technol. 7.

Nasiri, A., Zhao, Y., Gan, H., 2024b. Automated detection and counting of broiler behaviors using a video recognition system. Comput. Electron. Agric. 221.

Neethirajan, S., 2017. Recent advances in wearable sensors for animal health management. Sens. Bio- Sens. Res. 12, 15–29.

Neethirajan, S., 2025. Adapting a large-scale transformer model to decode chicken vocalizations: A non-invasive AI approach to poultry welfare. AI 6 (4), 65.

Niloofar, P., Francis, D.P., Lazarova-Molnar, S., Vulpe, A., Vochin, M.-C., Suciu, G., Balanescu, M., Anestis, V., Bartzanas, T., 2021. Data-driven decision support in livestock farming for improved animal health, welfare and greenhouse gas emissions: Overview and challenges. Comput. Electron. Agric. 190, 106406.

Noor, A., Zhao, Y., Koubâa, A., Wu, L., Khan, R., Abdalla, F.Y., 2020. Automated sheep facial expression classification using deep transfer learning. Comput. Electron. Agric. 175, 105528.

Norouzzadeh, M.S., Morris, D., Beery, S., Joshi, N., Jojic, N., Clune, J., 2021. A deep active learning system for species identification and counting in camera trap images. Methods Ecol. Evol. 12 (1), 150–161.

Nyamuryekungwe, S., Duff, G., Utsumi, S., Estell, R., McIntosh, M.M., Funk, M., Cox, A., Cao, H., Spiegal, S., Perea, A., Cibils, A.F., 2023. Real-time monitoring of grazing cattle using lora-wan sensors to improve precision in detecting animal welfare implications via daily distance walked metrics. Animals 13 (16).

Odena, A., Olah, C., Shlens, J., 2017. Conditional image synthesis with auxiliary classifier GANs. In: International Conference on Machine Learning. pp. 2642–2651.

Oliveira, D.A.B., Pereira, L.G.R., Bresolin, T., Ferreira, R.E.P., Dorea, J.R.R., 2021. A review of deep learning algorithms for computer vision systems in livestock. Livest. Sci. 253, 104700.

Palanisamy, V., Ratnarajah, N., 2021. Detection of wildlife animals using deep learning approaches: A systematic review. In: 2021 21st International Conference on Advances in ICT for Emerging Regions. ICter, pp. 153–158.

Pandiaraja, P., Madhumitha, U., Kumar, S.M., Santhosh, N., 2023. An analysis of wild fauna trespassing warning system using CNN and YOLO v3. In: 2023 3rd International Conference on Innovative Mechanisms for Industry Applications. ICIMIA, pp. 354–359.

Peng, J., Wang, D., Liao, X., Shao, Q., Sun, Z., Yue, H., Ye, H., 2020. Wild animal survey using UAS imagery and deep learning: Modified faster R-CNN for kiang detection in tibetan plateau. ISPRS J. Photogramm. Remote Sens. 169, 364–376.

Pereira, T.D., Tabris, N., Matsliah, A., Turner, D.M., Li, J., Ravindranath, S., Papadoyannis, E.S., Normand, E., Deutsch, D.S., Wang, Z.Y., et al., 2022. SLEAP: A deep learning system for multi-animal pose tracking. Nature Methods 19 (4), 486–495.

Petso, T., Jamisola, R.S., Mpoeleng, D., 2022. Review on methods used for wildlife species and individual identification. Eur. J. Wildl. Res. 68 (3).

Pillay, S.J., Bangira, T., Sibanda, M., Kebede Gurmessa, S., Clulow, A., Mabhaudhi, T., 2024. Assessing drone-based remote sensing for monitoring water temperature, suspended solids and CDOM in inland waters: A global systematic review of challenges and opportunities. Drones 8 (12), 733.

Prabhu, S., Sreenath, M.V., Malavika, V., Om, H., Swetha, S., 2023. Detection and recognition of animals using yolo algorithm. In: 2023 International Conference on Distributed Computing and Electrical Circuits and Electronics. ICDCECE, pp. 1–6.

Radford, A., Metz, L., Chintala, S., 2015. Unsupervised representation learning with deep convolutional generative adversarial networks. arXiv preprint arXiv:1511.06434.

Ravoor, P.C., Sudarshan, T.S.B., 2020. Deep learning methods for multi-species animal re-identification and tracking-a survey. Comput. Sci. Rev. 38, 100289.

Redmon, J., Divvala, S., Girshick, R., Farhadi, A., 2016. You only look once: Unified, real-time object detection. arXiv preprint arXiv:1506.02640.

Redmon, J., Farhadi, A., 2017. YOLO9000: Better, faster, stronger. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition.

Redmon, J., Farhadi, A., 2018. Yolov3: An incremental improvement. arXiv preprint arXiv:1804.02767.

Rencricea, G., Froldi, F., Moschini, M., Trevisan, M., Lamastra, L., 2023. Mitigation actions scenarios applied to the dairy farm management systems. Foods 12 (9), 1860.

Rickert, M., Klein, A., Adrion, F., Hoffmann, C., Gallmann, E., 2020. Automatically detecting pig position and posture by 2D camera imaging and deep learning. Comput. Electron. Agric. 174, 105391.

Robson, J.F., Denholm, S.J., Coffey, M., 2021. Automated processing and phenotype extraction of ovine medical images using a combined generative adversarial network and computer vision pipeline. Sensors 21 (21).

Rohan, A., Rafaq, M.S., Hasan, M.J., Asghar, F., Bashir, A.K., Dottorini, T., 2024. Application of deep learning for livestock behaviour recognition: A systematic literature review. Comput. Electron. Agric. 224, 109115.

Ronneberger, O., Fischer, P., Brox, T., 2015. U-net: Convolutional networks for biomedical image segmentation. In: Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015. Springer.

Roy, A.M., Bose, T.K., Raj, K., 2023. WilDect-YOLO: An efficient and robust computer vision-based accurate object localization model for automated endangered wildlife detection. Ecol. Inform. 75, 101919.

Sandler, M., Howard, A., Zhu, M., Zhmoginov, A., Chen, L.-C., 2018. Mobilenetv2: Inverted residuals and linear bottlenecks. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. pp. 4510–4520.

Schillings, J., Bennett, R., Rose, D.C., 2021. Exploring the potential of precision livestock farming technologies to help address farm animal welfare. Front. Anim. Sci. 2.

Schneider, S., Greenberg, S., Taylor, G.W., Kremer, S.C., 2020. Three critical factors affecting automated image species recognition performance for camera traps. Ecol. Evol. 10 (7), 3503–3517.

Selle, M., Spieß, F., Visscher, C., Rautenschlein, S., Jung, A., Auerbach, M., Hartung, J., Sürie, C., Distl, O., 2023. Real-time monitoring of animals and environment in broiler precision farming - how robust is the data quality? Sustainability 15 (21), 15527.

Seo, H., Ahn, D., Kim, S., Lee, Y., Chung, D., Park, S., 2020. EmbeddedPigDet fast and accurate pig detection for embedded board implementations. Appl. Sci. 10 (8), 2878.

Shang, C., Zhao, H., Wang, M.L., Wang, X.L., Jiang, Y., Gao, Q., 2023. Individual identification of cashmere goats via method of fusion of multiple optimization. Comput. Animat. Virtual Worlds 34 (2), e2048.

Sharma, T., Cline, D.E., Edgington, D., 2024a. Making use of unlabeled data: Comparing strategies for marine animal detection in long-tailed datasets using self-supervised and semi-supervised pre-training. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 1224–1233.

Sharma, A., Kumar, V., Longchamps, L., 2024b. Comparative performance of YOLOv8, YOLOv9, YOLOv10, YOLOv11 and faster R-CNN models for detection of multiple weed species. Smart Agric. Technol. 9, 100648.

Shulkin, J.D., Vibhatasilpin, A., Adhana, V., 2025. Pervasive sensing for livestock health and activity monitoring: Current methods and techniques. arXiv preprint arXiv:2503.15674.

Simonyan, K., Zisserman, A., 2014. Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556.

Singh, P., Devi, K.J., Varish, N., 2021. Muzzle pattern based cattle identification using generative adversarial networks. In: Soft Computing for Problem Solving. In: Advances in Intelligent Systems and Computing, vol. 1392, Springer, Singapore, pp. 13–23.

Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V., Rabinovich, A., 2014. Going deeper with convolutions. arXiv preprint arXiv:1409.4842.

Taiwo, G., Vadera, S., Alameer, A., 2025. Vision transformers for automated detection of pig interactions in groups. Smart Agric. Technol. 10, 100774.

Tan, M., Le, Q., 2019. Efficientnet: Rethinking model scaling for convolutional neural networks. In: International Conference on Machine Learning. PMLR.

Tan, D., Liang, X., 2023. Multiclass malaria parasite recognition based on transformer models and a generative adversarial network. Sci. Rep. 13 (1), 17136.

Tassinari, P., Bovo, M., Benni, S., Franzoni, S., Poggi, M., Mammi, L.M.E., Mattoccia, S., Di Stefano, L., Bonora, F., Barbaresi, A., et al., 2021. A computer vision approach based on deep learning for the detection of dairy cows in free stall barn. Comput. Electron. Agric. 182, 106030.

Tjaudrasuwita, M., Sun, J.J., Kennedy, A., Chaudhuri, S., Yue, Y., 2021. Interpreting expert annotation differences in animal behavior. arXiv preprint arXiv:2106.06114.

Tu, S., Zeng, Q., Liu, H., Liang, Y., Liu, X., Huang, L., Huang, Z., 2023. MaskDis R-CNN: An instance segmentation algorithm with adversarial network for herd pigs. IET Image Process. 17 (12), 3488–3499.

Tuia, D., Kellenberger, B., Beery, S., Costelloe, B.R., Zuffi, S., Risse, B., Mathis, A., Mathis, M.W., Van Langevelde, F., Burghardt, T., et al., 2022. Perspectives in machine learning for wildlife conservation. Nat. Commun. 13 (1).

Valletta, J.J., Torney, C., Kings, M., Thornton, A., Madden, J., 2017. Applications of machine learning in animal behaviour studies. Anim. Behav. 124, 203–220.

Varghese, R., M, S, 2024. Yolov8: A novel object detection algorithm with enhanced performance and robustness. In: 2024 International Conference on Advances in Data Engineering and Intelligent Computing Systems. ADICS, pp. 1–6.
Von Ziegler, L., Sturman, O., Bohacek, J., 2021. Big behavior: Challenges and opportunities in a new era of deep behavior profiling. Neuropsychopharmacology 46 (1), 33–44.
Wäldchen, J., Mäder, P., 2018. Machine learning for image-based species identification. Methods Ecol. Evol. 9 (11), 2216–2225.
Walter, T., Couzin, I.D., 2021. TRex, a fast multi-animal tracking system with markerless identification, and 2D estimation of posture and visual fields. ELife 10, e64000.
Wang, A., Chen, H., Lin, L., Chen, K., Lin, Z., Han, J., et al., 2024a. Yolov10: Real-time end-to-end object detection. Adv. Neural Inf. Process. Syst. 37, 107984–108011.
Wang, C., Gao, G., Wang, J., Lv, Y., Li, Q., Li, Z., Zhang, X., Wu, H., 2023a. Getvae-gan: An image enhancement network for low-light cattle farm scenes by integrating fusion gate transformation mechanism and variational autoencoder GAN. IEEE Access 11, 126650–126660.
Wang, Y., Kang, X., Chu, M., Liu, G., 2022. Deep learning-based automatic dairy cow ocular surface temperature detection from thermal images. Comput. Electron. Agric. 202, 107429.
Wang, Z., Shadpour, S., Chan, E., Rotondo, V., Wood, K.M., Tulpan, D., 2021. ASAS-NANP SYMPOSIUM: Applications of machine learning for livestock body weight prediction from digital images. J. Anim. Sci. 99 (2), skab022.
Wang, C.Y., Yeh, I.H., Liao, H.Y.M., 2024b. YOLOv9: Learning what you want to learn using programmable gradient information. arXiv preprint arXiv:2402.12345.
Wei, Y.-C., Chen, W.-L., Tuanmu, M.-N., Lu, S.-S., Shiao, M.-T., 2024. Advanced montane bird monitoring using self-supervised learning and transformer on passive acoustic data. Ecol. Inform. 84, 102927.
Wu, Z., Zhang, T., Fang, C., Yang, J., Ma, C., Zheng, H., Zhao, H., 2023. Super-resolution fusion optimization for poultry detection: A multi-object chicken detection method. J. Anim. Sci. 101.
Xu, B., Wang, W., Falzon, G., Kwan, P., Guo, L., Chen, G., Tait, A., Schneider, D., 2020a. Automated cattle counting using mask R-CNN in quadcopter vision system. Comput. Electron. Agric. 171, 105300.
Xu, B., Wang, W., Falzon, G., Kwan, P., Guo, L., Sun, Z., Li, C., 2020b. Livestock classification and counting in quadcopter aerial images using mask R-CNN. Int. J. Remote Sens. 41 (21), 8121–8142.
Yari, Y., Naeve, I., Hammerdal, A., Bergtun, P.H., Masoy, S.-E., Voormolen, M.M., Løvstakken, L., 2024. Automated measurement of ovary development in atlantic salmon using deep learning. Ultrasound Med. Biol. 50 (3), 364–373.
Yogi, M.K., Prasad, P.S., Swetha, C.S.S., Lakshmi, K.N.S., 2024. Deploying XAI with IoT for the protection of endangered species. In: Green Industrial Applications of Artificial Intelligence and Internet of Things. Bentham Science Publishers, pp. 183–198.
Yousefi, D.M., Rafie, A.M., Al-Haddad, S., Azrad, S., 2022. A systematic literature review on the use of deep learning in precision livestock detection and localization using unmanned aerial vehicles. IEEE Access 10, 80071–80091.
Yu, Z., Wan, L., Yousaf, K., Lin, H., Zhang, J., Jiao, H., Yan, G., Song, Z., Tian, F., 2024. An enhancement algorithm for head characteristics of caged chickens detection based on cyclic consistent migration neural network. Poult. Sci. 103 (6).
Yu, H., Xu, Y., Zhang, J., Zhao, W., Guan, Z., Tao, D., 2021. AP-10k: A benchmark for animal pose estimation in the wild. arXiv preprint arXiv:2108.12617.
Zhai, Q., Li, X., Yang, F., Chen, C., Cheng, H., Fan, D.-P., 2021. Mutual graph learning for camouflaged object detection. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 12997–13007.
Zhang, Q., Ahmed, K., Sharda, N., Wang, H., 2024a. A comprehensive survey of animal identification: Exploring data sources, AI advances, classification obstacles and the role of taxonomy. Int. J. Intell. Syst. 2024 (1), 7033535.
Zhang, K., Han, S., Wu, J., Cheng, G., Wang, Y., Wu, S., Liu, J., 2023. Early lameness detection in dairy cattle based on wearable gait analysis using semi-supervised LSTM-autoencoder. Comput. Electron. Agric. 213, 108252.
Zhang, J., Yang, G., Sun, L., Zhou, C., Zhou, X., Li, Q., Bi, M., Guo, J., 2021. Shrimp egg counting with fully convolutional regression network and generative adversarial network. Aquac. Eng. 94, 102175.
Zhang, S., Yang, X., Wang, Y., Zhao, Z., Liu, J., Liu, Y., Sun, C., Zhou, C., 2020. Automatic fish population counting by machine vision and a hybrid deep neural network model. Animals 10 (2), 364.
Zhang, X., Zhu, R., Zheng, W., Chen, C., 2024b. Detection of leg diseases in broiler chickens based on improved YOLOv8 X-ray images. IEEE Access 12, 47385–47401.
Zheng, X., Li, F., Lin, B., Xie, D., Liu, Y., Jiang, K., Gong, X., Jiang, H., Peng, R., Duan, X., 2022. A two-stage method to detect the sex ratio of hemp ducks based on object detection and classification networks. Animals 12, 1177.
Zhu, J.-Y., Park, T., Isola, P., Efros, A.A., 2017. Unpaired image-to-image translation using cycle-consistent adversarial networks. In: Proceedings of the IEEE International Conference on Computer Vision. ICCV.
Zhuge, M., Lu, X., Guo, Y., Cai, Z., Chen, S., 2022. Cubenet: X-shape connection for camouflaged object detection. Pattern Recognit. 127 (C).