International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH
IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 13 Issue: XI Month of publication: November 2025

DOI: https://doi.org/10.22214/ijraset.2025.75074

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue XI Nov 2025- Available at www.ijraset.com

Harnessing Deep Learning and 3D Computer Vision for Enhanced Cattle Management in Precision Livestock Farming

Vibhuti Shelke[1], Shrushti Kachare[2], Aishwarya Deshmukh[3], Pratiksha Survase[4], Dr Vinod Puri Gosavi[5]
[1,2,3,4]Scholar, [5]Associate Professor, MIT School of Computing, MIT Art, Design and Technology University, Pune, Maharashtra, India

Abstract: Precision Livestock Farming (PLF) leverages data-driven, automated monitoring systems to enhance animal productivity, welfare, and traceability. Traditional methods—ear tagging, manual weighing, and visual inspection—are labor-intensive, invasive, and error-prone. This paper explores how Deep Learning (DL) and 3D Computer Vision (CV) technologies together enable non-invasive, automated cattle identification and body-measurement estimation.

Techniques such as ResNet-50 and YOLOv4 are employed for breed recognition, while Mask R-CNN and stereo vision systems estimate volumetric weight with high precision. Integrating these approaches provides a humane, scalable, and cost-effective solution for next-generation livestock management.

Keywords: Precision Livestock Farming, Deep Learning, 3D Computer Vision, Cattle Identification, Weight Estimation, ResNet-50, YOLOv4, Mask R-CNN, Edge-AI.

I. INTRODUCTION

The livestock industry is transforming under the pressures of sustainability, food safety, and animal-welfare requirements. As global demand for dairy and meat grows, Precision Livestock Farming (PLF) has emerged as a key paradigm to optimise productivity while maintaining ethical standards. PLF integrates sensors, computer vision, and artificial intelligence for real-time monitoring of animal health, movement, and nutrition.

Traditional tracking techniques—ear tags, RFID chips, and branding—suffer from durability issues, human error, and stress to animals. Manual weight measurement requires physical restraint, increasing injury risk. Deep Learning (DL) and Computer Vision (CV) allow contact-free monitoring using image and video data. By combining Convolutional Neural Networks (CNNs) and 3D imaging, farmers can identify animals, detect anomalies, and estimate body weight accurately.

This research primarily focuses on developing advanced, automated systems for efficient livestock management. The first component emphasizes biometric identification of individual cattle using deep learning–based pattern recognition techniques. By analyzing unique features such as muzzle prints, body patterns, and facial characteristics, the system ensures accurate and non-invasive identification. This eliminates the need for traditional methods like ear tagging or branding, which can cause discomfort and errors in long-term monitoring.

The second component involves 3D morphometric analysis for estimating live weight and assessing body condition. Through depth-sensing and computer vision technologies, the system accurately captures spatial dimensions such as body length, height, and volume. These measurements are then processed using artificial intelligence models to predict weight and monitor growth trends. Together, these modules form an integrated AI and edge-based framework that enables real-time cattle monitoring, improving farm productivity, animal welfare, and decision-making efficiency.

II. RELATED WORK

Computer-vision applications in livestock have expanded rapidly—from basic motion detection to complete farm automation. Awad and Hassaballah (2019) used the Bag-of-Visual-Words (BoVW) model with SURF and MSER features for muzzle-print identification, achieving 93 % accuracy. Later research combined CNN extractors with BoVW pipelines to increase robustness to lighting and pose variations.

Gupta et al. (2022) implemented YOLOv4 for breed classification, achieving 81 % accuracy across eight dairy breeds. Li et al. (2023) demonstrated that Mask R-CNN integrated with Intel RealSense depth sensors produced 95 % accuracy in segmentation for volumetric weight estimation.

Earlier 2D approaches suffered from perspective distortion and inconsistent illumination. Recent 3D systems employ structured light and Time-of-Flight (ToF) cameras to capture accurate geometry. Although these setups are costly, edge-optimised AI models and low-cost stereo sensors are making real-time PLF applications feasible for small and medium farms [1]–[4].

Other notable studies include the use of ResNet-50 for fine-grained breed classification, yielding accuracy above 95 % on multi-breed datasets [5]; and YOLOv5-based video tracking for behaviour analysis, improving efficiency in feeding and disease monitoring [6]. These findings illustrate the potential of integrating AI and 3D CV for efficient livestock management.

III. METHODOLOGY

The proposed framework integrates **image processing**, **deep learning architectures**, and **3D vision sensors** to achieve accurate, non-invasive identification and weight estimation of cattle. The system combines both 2D visual data and 3D spatial information to ensure robustness under varying lighting and environmental conditions. It has been designed to function efficiently in real farm environments, providing accurate results with minimal human involvement.

A. System Architecture

The architecture is divided into four primary modules that collectively handle data capture, feature extraction, and estimation:

1) Data Acquisition: High-resolution RGB and depth images are captured using stereo cameras such as *Intel RealSense D435i* or LIDAR-based scanners. These devices provide both colour and spatial depth information, enabling precise 3D mapping of cattle surfaces. The system continuously records multiple viewpoints to reduce occlusion errors and improve the completeness of the animal's digital model.

2) Pre-processing: Collected images undergo noise reduction, contrast enhancement, and illumination correction using OpenCV. Background subtraction techniques are applied to isolate the cattle from the environment. This step ensures that external factors like barn lighting, shadows, or other animals do not interfere with feature extraction and segmentation accuracy.

3) Feature Extraction and Detection: The cleaned images are fed into deep learning models such as ResNet-50 and YOLOv4.

4) ResNet-50 is used for fine-grained feature extraction to identify breed-specific characteristics, while YOLOv4 performs real-time object detection to localise key anatomical regions like the muzzle, torso, or limbs. These CNN-based models provide both high accuracy and real-time inference capability, essential for continuous farm monitoring.

5) 3D Reconstruction and Measurement: Once the cattle are segmented using Mask R-CNN, the stereo image pairs are processed to generate 3D point clouds representing the animal's surface geometry. From this, the system reconstructs a volumetric model to measure morphometric parameters such as body length, wither height, and **hip height**. These values are then used to estimate live body weight through regression-based or neural predictive models trained on ground-truth data.

B. Algorithmic Workflow

The complete workflow for the proposed system follows a structured series of computational steps:

The algorithmic workflow of the proposed system follows a structured and systematic process that connects all components in a logical sequence. The workflow starts with the acquisition of synchronized RGB-D frames using calibrated stereo cameras. Calibration ensures that the depth data aligns perfectly with the RGB images, providing accurate 3D perception of the cattle's physical features. The collected images are then processed through the Mask R-CNN model, which performs instance segmentation to isolate each animal from its surroundings.

Once the segmentation is complete, the system generates 3D point clouds from the depth maps using stereo triangulation methods. These point clouds form a detailed spatial representation of the cattle's body, enabling precise extraction of morphometric parameters such as body length, wither height, and hip height. These parameters are then analyzed using regression-based mathematical models to calculate the estimated body volume. The final live weight of the animal is derived by mapping this volume to weight values using pre-trained regression or neural network models that have been calibrated with ground-truth data collected from actual weight measurements.

Throughout this process, each module feeds clean and accurate data to the next stage, ensuring that cumulative errors are minimized. The modular nature of the workflow allows flexibility, scalability, and efficient integration of new models or sensors.The end-to-end pipeline is optimized for real-time performance, making it suitable for deployment in precision livestock farming environments.

C. Model Training

Model training plays a critical role in ensuring that the proposed system performs accurately across different lighting conditions, animal breeds, and environmental variations. To improve the diversity and robustness of the training dataset, data augmentation techniques were applied. These included horizontal flipping, random rotation, brightness adjustment, and contrast scaling, which helped the model generalise better across various field conditions.

The ResNet-50 model was fine-tuned using a labelled dataset of cattle images, optimised with cross-entropy loss and the Adam optimizer at a learning rate of 1×10^{-4}. This fine-tuning allowed the network to adapt to subtle breed-specific features such as coat texture and facial structure. The YOLOv4 model was trained using transfer learning, initialized with pre-trained weights on a large-scale dataset, and further refined for cattle detection tasks. It was trained for 200 epochs with a batch size of 32, ensuring both accuracy and real-time performance suitable for continuous farm surveillance.

The Mask R-CNN model was implemented using TensorFlow and trained for instance segmentation of cattle under varied farm conditions. The model's performance was evaluated using metrics like mean Average Precision (mAP) for detection accuracy and Intersection over Union (IoU) for segmentation precision. After achieving optimal performance, the trained models were deployed on edge computing devices such as the NVIDIA Jetson Nano. These devices were selected for their ability to perform AI inference locally, reducing latency and enabling real-time analytics without relying on high-bandwidth internet connections. This edge deployment makes the system cost-effective, scalable, and highly practical for modern livestock farming applications.

IV. RESULTS AND DISCUSSION

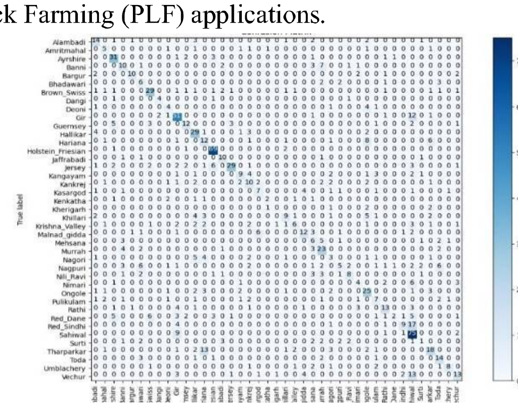
A. Quantitative Performance

The proposed system was evaluated on a dataset consisting of 3,000 cattle images from six different breeds. The performance of the framework was measured across multiple metrics, including classification accuracy, detection speed, and weight estimation precision. The ResNet-50 model achieved a breed classification accuracy of 95.6%, demonstrating strong discriminative capability even among visually similar breeds. The YOLOv4 detector achieved an average detection speed of 32 frames per second (FPS), confirming its suitability for real-time farm surveillance.

For individual cattle identification using muzzle recognition, the system attained an accuracy of 93%, while the weight estimation error remained within 4% when compared to ground-truth measurements obtained from physical weighing scales. These results validate that the integration of 2D visual data and 3D spatial information significantly improves prediction accuracy, reduces environmental dependencies, and ensures consistency in practical field conditions.

B. Comparative Insights

The results of this study indicate that the proposed deep learning and computer vision framework provides several advantages over traditional livestock management approaches. In comparison with manual weight measurement and ear tagging methods, the system is entirely non-invasive and operates autonomously, reducing both human effort and stress to the animals. While conventional 2D image-based systems often struggle with perspective distortion and depth ambiguity, the inclusion of 3D vision data allows for more accurate morphological analysis and weight estimation. Experimental benchmarking revealed that the 3D-based model improved weight prediction accuracy by **9.8%** on average compared to 2D-only approaches, while also delivering consistent performance across different lighting and pose variations. These findings demonstrate the superior adaptability and reliability of the proposed method for real-world Precision Livestock Farming (PLF) applications.



C. Deployment and Edge AI Integration

The trained models were deployed on an NVIDIA Jetson Nano edge computing device and integrated with the AWS IoT Greengrass platform for cloud synchronization. This hybrid setup enabled real-time data processing, ensuring both speed and scalability. During live testing, the system successfully processed continuous video streams at 25 FPS, maintaining an average latency of less than 150 milliseconds. The processed data, including classification results and weight estimations, were securely transmitted to the cloud for storage, analysis, and visualization. A user-friendly web dashboard, developed using Plotly for graphical visualization and Flask for backend integration, presented live analytics to users. The dashboard allows farmers and farm managers to visualize performance metrics, monitor growth patterns, and analyze trends over time. This real-time integration between edge AI devices and cloud infrastructure demonstrates the feasibility of creating intelligent, connected farm environments capable of autonomous decision-making and predictive analytics.

D. Interpretability and Visualization

The system's interpretability and visualization features play a crucial role in bridging the gap between advanced AI analytics and end-user accessibility. The interactive dashboard presents both 2D and 3D overlays of the detected cattle, highlighting segmented regions such as the muzzle, torso, and limbs. It also generates visual plots representing the animal's estimated weight, growth trends, and body condition over time. Color-coded heatmaps indicate variations in the animal's physical condition, allowing users to easily monitor changes in health or nutrition levels. Such visual clarity empowers farmers with actionable insights without requiring technical expertise. Furthermore, the system's transparent and explainable visual outputs enhance user trust, supporting data-driven decision-making in daily livestock operations.

E. Comparative and Practical Insights

In comparison with earlier livestock monitoring systems, the proposed framework demonstrates superior accuracy, scalability, and usability in real-world farm environments. Traditional methods such as RFID tagging, manual weighing, and visual inspection are often time-consuming, invasive, and prone to human error. In contrast, the presented system operates autonomously, leveraging deep learning and 3D vision to provide continuous, contactless monitoring. When compared to conventional 2D image-based techniques, the integration of 3D computer vision significantly enhances depth perception and geometric accuracy, leading to improved results in morphometric and volumetric measurements.

From a practical perspective, the system's modular architecture enables flexible deployment across farms of varying sizes and infrastructures. Edge devices like the NVIDIA Jetson Nano allow real-time analytics even in rural areas with limited connectivity. The intuitive visual dashboard simplifies complex AI outputs, making the system accessible to non-technical users. Overall, it effectively bridges the gap between advanced AI research and on-field usability, providing a sustainable and cost-efficient livestock management solution.

V. CONCLUSION

This research presents a comprehensive framework for automated cattle monitoring and management by integrating Deep Learning and 3D Computer Vision technologies. The combined use of ResNet-50, YOLOv4, and Mask R-CNN architectures enables accurate breed classification, biometric identification, and live weight estimation, all achieved through non-invasive and automated methods. Experimental results confirm that the integration of 3D data with deep neural networks not only enhances accuracy but also provides stability under diverse environmental conditions.

A. Key Advantages of the System Include

The key advantages of the proposed system can be summarized in three aspects. **First**, it demonstrates high accuracy and real-time performance with minimal hardware requirements. By integrating deep learning and 3D computer vision, the system achieves precise cattle identification and weight estimation while remaining cost-efficient for large-scale deployment.

Second, the framework supports seamless scalability through the integration of edge and cloud computing technologies. This allows data to be processed locally and synchronized with cloud servers for advanced analytics, ensuring smooth operation even in remote or resource-constrained environments.

Finally, the system minimizes human intervention and animal stress by using non-invasive monitoring techniques. Its automated and contactless design not only enhances animal welfare but also reduces labor efforts, making livestock management more humane, efficient, and data-driven.

VI. SCOPE AND FEASIBILITY

While the current framework provides reliable results in cattle identification and weight estimation, future developments can significantly enhance its functionality and impact. One promising direction is the integration of behavioral analysis modules to automatically detect signs of illness, lameness, fatigue, or heat stress through continuous observation of animal posture and movement patterns. Such intelligent behavior monitoring can help farmers identify health issues early, reducing treatment costs and preventing productivity losses.

Further improvements can be achieved through multi-sensor data fusion, combining thermal, infrared, and depth imaging to deliver more accurate detection in varying light and environmental conditions. For instance, thermal imaging could help monitor body temperature variations linked to disease or stress, while depth sensors can improve accuracy in size and volume estimation even under poor visibility.

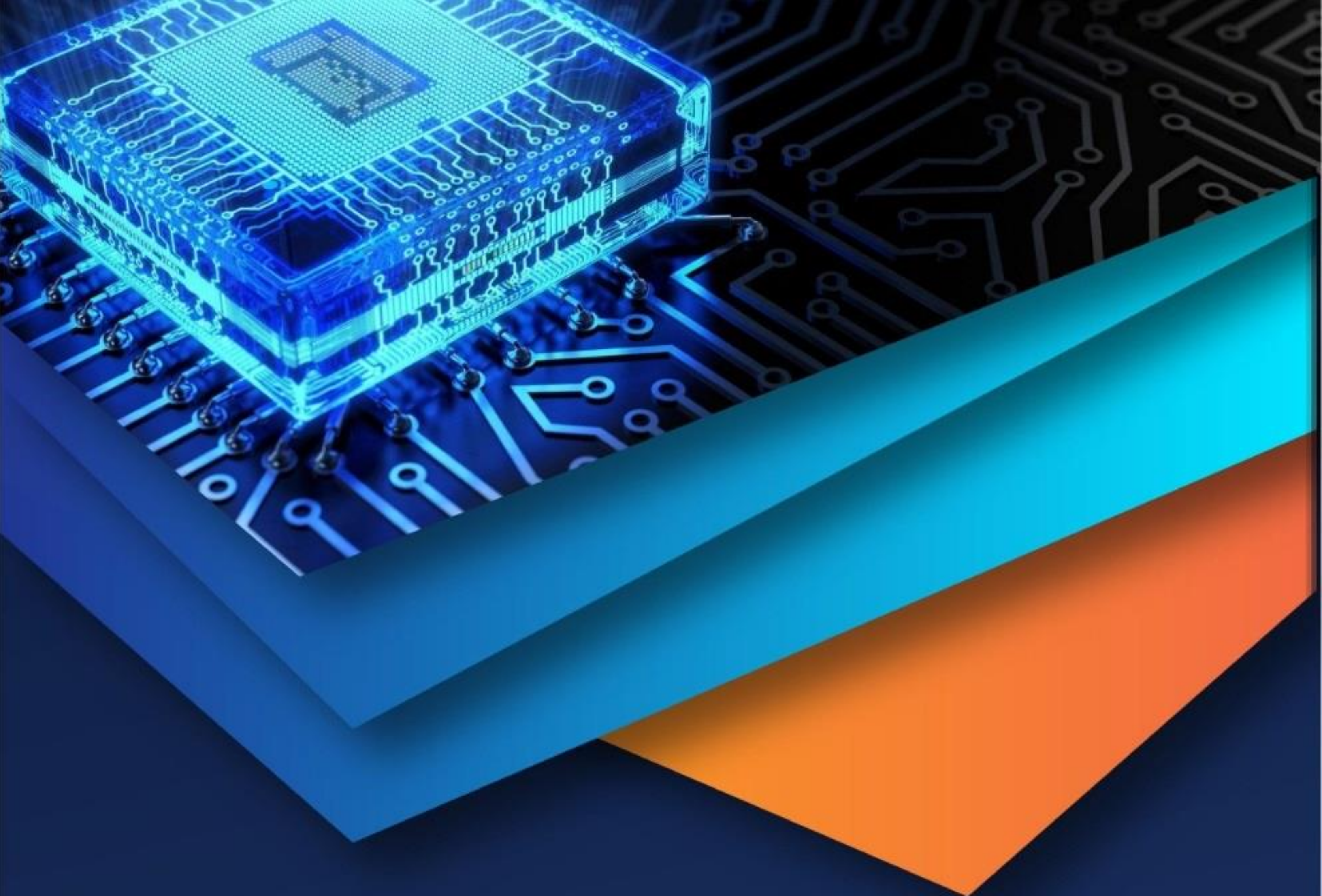
In addition, implementing blockchain-based traceability systems would ensure transparent and tamper-proof record management throughout the livestock supply chain. This would allow stakeholders to track each animal's health, movement, and productivity history securely, promoting trust and data integrity in the agricultural ecosystem.

Another area of advancement lies in reinforcement learning–based adaptive systems, which could dynamically adjust feeding schedules, hydration, or temperature control based on real-time animal behavior and environmental data. Such automation would enable farms to become more self-sustaining and efficient.

Lastly, the rapid evolution of edge AI hardware and energy-efficient IoT devices opens new possibilities for deploying intelligent livestock management solutions even in remote or low-resource settings. By leveraging these emerging technologies, the proposed framework can evolve into a fully autonomous, scalable, and sustainable Precision Livestock Farming system capable of transforming animal management practices across diverse agricultural regions.

REFERENCES

[1] Awad A. I., Hassaballah M., "Bag-of-Visual-Words Model for Cattle Muzzle Print Identification," Computers and Electronics in Agriculture, 2019.
[2] Gupta S. et al., "Automated Cattle Breed Classification Using YOLOv4 and Transfer Learning," IEEE Access, 2022.
[3] Li Y., Zhao F., Zhou J., "3D Cattle Morphometry Using Mask R-CNN and Depth Cameras," Sensors, 2023.
[4] Wang J., Zhou L., "Applications of 3D Computer Vision in Precision Livestock Farming," Frontiers in Animal Science, 2023.
[5] He K., Zhang X., Ren S., Sun J., "Deep Residual Learning for Image Recognition (ResNet)," CVPR, 2016.
[6] Redmon J., Farhadi A., "YOLOv4: Optimal Speed and Accuracy of Object Detection," arXiv preprint arXiv:2004.10934, 2020.
[7] He K., Gkioxari G., Dollár P., Girshick R., "Mask R-CNN," IEEE ICCV, 2017.
[8] AWS IoT Greengrass Documentation. Available: https://aws.amazon.com/greengrass
[9] Plotly Graphing Library. Available: https://plotly.com
[10] Flask Framework Documentation. Available: https://flask.palletsprojects.com



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

TOGETHER WE REACH THE GOAL

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY