ISSN No. 0976-5697

VOLUME 17 SPECIAL ISSUE 1, JULY 2026
International Journal of Advanced Research in Computer Science

RESEARCH PAPER

Available Online at www.ijarcs.info

DEEP LEARNING-DRIVEN CATTLE AND BUFFALO CLASSIFICATION USING CNN- YOLO HYBRID FRAMEWORK FOR SMART LIVESTOCK MANAGEMENT

Ms. Priya Kumawat, Divisha Upadhyay, Chiraushi Dadhich
Department of Computer Science & Engineering
Geetanjali Institute of Technical Studies
Udaipur, India
kumawatpriya21@gmail.com, divishaaaupadhyay@gmail.com, chiraushi@gmail.com

Abstract—Accurate identification and classification of livestock animals is a fundamental requirement in modern agricultural management. This paper proposes an image- based deep learning framework to automatically distinguish between cattle and buffaloes using Convolutional Neural Networks (CNN) integrated with a YOLO-based object detection pipeline. A curated dataset of 2,500 annotated images— comprising 1,300 cattle and 1,200 buffalo images— was utilized for training and evaluation. The preprocessing pipeline applies resizing to 224×224 pixels, normalization, and data augmentation techniques including random rotation, horizontal flipping, and brightness adjustment to enhance model generalization. The CNN model extracts multi-scale hierarchical features, while YOLO localizes individual animals within cluttered farm backgrounds prior to classification. Experimental evaluation demonstrates an overall classification accuracy of 93.2 %, with precision of 0.92 and recall of 0.94. The proposed system significantly reduces dependence on manual observation and supports real-time, automated livestock monitoring. These results confirm the viability of deep learning for scalable and accurate animal type classification in smart farming environments.

Keywords—Deep Learning, Convolutional Neural Network, YOLO, Image Classification, Cattle, Buffalo, Smart Agriculture, Transfer Learning.

I. INTRODAUCTION

Livestock farming constitutes a fundamental pillar of the global agricultural economy, supporting billions of livelihoods and contributing significantly to food security. In countries such as India, cattle and buffaloes represent the backbone of the dairy sector, draft power, and rural income. Accurate identification and classification of these animals are essential for breed management, veterinary monitoring, insurance processing, productivity assessment, and livestock traceability in large-scale farm operations. Traditional animal identification techniques rely on physical tagging, branding, and expert visual inspection. These approachesare not only labor-intensive and time-consuming but also inherently subjective and prone to human error. When implemented at large scale, inconsistencies in manual assessment can lead to erroneous breed records, poor breeding decisions, and mismanagement of livestock health programs

The rapid advancement of Artificial Intelligence (AI) and computer vision has opened transformative avenues for automating complex morphological traits Cattle and buffaloes can appear visually similar to untrained observers under variable lighting, cluttered backgrounds, and diverse poses.

Distinguishing between these two species requires automated extraction of subtle morphological cues—such as hump structure, horn curvature, body profile, and ear orientation—that are reliably encoded in high- quality digital images.

II. LITERATURE REVIEW

Livestock evaluation forms the foundation of the dairy industry, directly influencing decisions related to breeding, nutrition, productivity, and economic management. Animal Type Classification (ATC) plays a central role in determining phenotypic characteristics of animals such as cattle and buffaloes. Accurate assessment of body conformation, muscle structure, horn shape, skin texture, and hump formation is essential to predict milk yield potential, adaptability, reproductive efficiency, and overall longevity Traditionally, ATC is performed by trained experts through visual inspection and manual measurement of specific body parts. While this method has served the livestock sector for decades, it suffers from several drawbacks. The process is time-consuming, labour- intensive, and highly dependent on human expertise and judgement. Different evaluators may produce inconsistent results due to subjectivity, fatigue, or environmental distractions such as poor lighting,animal movement,or background clutter.

Consequently,data obtained through manual evaluation often lacks standardization and repeatability, decade. Early approaches relied on handcrafted features such as Local Binary Patterns (LBP), Histogram of Oriented Gradients (HOG), and Gabor filters.

The introduction of deep CNNs transformed the field fundamentally. Krizhevsky et al. demonstrated the power of deep convolutional architectures through AlexNet, which achieved unprecedented accuracy on the ImageNet benchmark

III. OBJECTIVE OF THE STUDY

The primary objective of this research is to design,develop, and validate an image-based classification framework capable of accurately distinguishing between cattle and buffaloes using deep learning. The specific objectives are as follows:

(1) Curate a balanced and diverse annotated image dataset representative of real-world farm conditions.

(2) Implement an effective preprocessing and augmentation pipeline to improve model robustness.

(3) Develop a CNN-based feature extraction and classification model achieving accuracy ≥ 90 %.

(4) Integrate a YOLO-based detection stage to localise individual animals prior to classification

(5) Evaluate model performance using accuracy, precision, recall, F1-score, and confusion matrix metrics.

(6) Demonstrate feasibility of deployment on mobile or embedded platforms for practical smart farming use.

IV. PROPOSED METHODOLOGY

The proposed work on image-based Animal Type Classification for Cattle and Buffaloes aims to develop an automated system that can identify and classify animals using deep learning and computer vision. The system follows a structured methodology that includes several stages: data acquisition, preprocessing, feature extraction, model development, training and testing, evaluation, and deployment.

A. *Image Preprocessing*

Before feeding images into the model, they undergo a preprocessing stage to enhance quality and standardise input formats. Preprocessing involves:

- Resizing all images to a fixed resolution (224×224 pixels) suitable for CNN input.
- Noise removal using Gaussian or median filters to eliminate background disturbances.
- Normalisation of pixel values to the range [0, 1] to improve convergence during training.

Contrast enhancement and histogram equalisation to ensure consistent image brightnesse.

This step ensures that all images are consistent and suitable for features are extraction by the neural network.

B. *Data Augmentation*

Data augmentation techniques are applied to improve generalisation and address dataset imbalance. Augmentation operations include: random horizontal flipping ($p = 0.5$); random rotation within ±20°;random brightness and contrast jitter in the range [0.8, 1.2]; random scaling and cropping; and Gaussian noise injection ($\sigma = 0.01$). These augmentations simulate real-world field variations and improve model robustness to pose, illumination, and background differences.

C. *CNN Architecture*

Feature extraction is the process of identifying unique characteristics that help distinguish cattle from buffaloes. Unlike traditional machine learning methods that rely on handcrafted features, deep learning models automatically learn discriminative visual patterns from images. Using a CNN, the system extracts hierarchical features at three levels-

- Low-level features: edges, corners, and textures.
- Mid-level features: body shapes, contours, and colour gradients.
- High-level features: distinct physical traits such as hump (present in cattle), horn shape (curved in buffaloes), skin tone, and ear orientation.

D. *Training Process*

A deep CNN architecture such as VGG16, ResNet- 50, or a custom-built CNN is used for classification. The model is trained using supervised learning, where labelled images are used to adjust network parameters. The Adam optimiser is employed for adaptive learning rate optimisation, and the binary cross-entropy loss function is used to measure prediction error. The model was trained for 50 epochs with a batch size of 32 on GPU hardware.

Early stopping (patience = 10 epochs) and dropout regularisation ($p = 0.5$) were applied to prevent overfitting.

V. MATHEMATICAL MODEL

A. *Convolution Operation*

convolution. Let the input image be represented as $I(x, y)$. The convolution operation is defined as:

$$F(i,j)=\Sigma_m\Sigma_n I(m,n)K(i-m,jn) \quad (1)$$

where $F(i, j)$ denotes the output activation at position (i, j), $I(m, n)$ is the input pixel value, and K is the learnable convolutional kernel.

Applying a Rectified Linear Unit (ReLU) activation introduces non-linearity: $A(i, j) = \max(0, F(i, j))$.

B. *Softmax Classification*

The final classification probabilities are computed using the Softmax function over the output logit vector z:

$$P(\text{class}_i) = e^{(z_i)} / \Sigma_j e^{(z_j)} \quad (2)$$

where z_i is the logit corresponding to class i. For the binary case (Cattle, Buffalo), the class with the higher posterior probability constitutes the predicted label.

C. *Binary Cross-Entrophy Loss*

$$L = -(1/N)\ \Sigma_i [\ y_i \log(\hat{p}_i) + (1 - y_i) \log(1 - \hat{p}_i)\ ] \quad (3)$$

where N is the number of training samples, y_i is the ground truth label, and $\hat{p}_i$ is the predicted probability for the positive class.

VI. SYSTEM ARCHITECTURE

The end-to-end system pipeline consists of the following sequential stages:

(1) Image Acquisition: Raw images are captured via farm surveillance cameras, smartphones, or uploaded through a graphical user interface (GUI).

(2) Preprocessing Module: Images are resized, normalised, and denoised.

(3) YOLO Detection Stage: The preprocessed frame is passed to YOLOv5, which outputs bounding boxes around detected animals. Each detected region is cropped and forwarded individually to the classifier [15].

(4) CNN Feature Extraction and Classification: The cropped ROI is input to fine- tuned ResNet-50. Hierarchical feature maps are computed across residual blocks, culminating in a flattened feature vector fed to the fully connected classification head [3].

(5) Softmax Output: The classification head outputs probability scores. The class with maximum probability (Cattle or Buffalo) is returned as the final prediction.

(6) Result Logging: The predicted label and confidence score are displayed to the operator and optionally logged to a livestock management database for record keeping. For example: > Input: Image of a black animal with curved horns→ Output: Buffalo (0.94 probability).

VII. APPICATION IN SMART FARMING

This automated classification system offers several advantages in smart farming, veterinary record management, and animal health monitoring. It can be integrated with IoT-based surveillance cameras and mobile applications to provide real-time classification and tracking of livestock.

Furthermore, it supports efficient breed management, disease identification, and productivity analysis in large- scale farms Key application areas include:

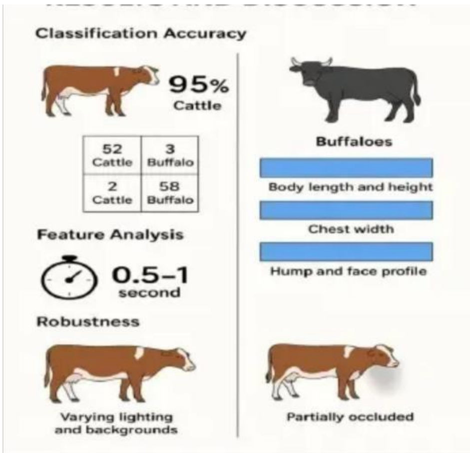
1. Automated animal inventory via farm surveillance cameras;
2. veterinary record management with automatic species logging; insurance and traceability verification for livestock registration programmes;
3. IoT and edge-device deployment using TensorFlow Lite or ONNX for real-time monitoring and breeding programme support through reliable phenotypic documentation.

AI driven systems align with the border goal of precision

precision agriculture and smart farming ,where technology is used to optimized agriculture and smart productivity while minimising labour and resource usage.

VIII. CONCLUSION

Image-based Animal Type Classification for cattle and buffaloes provides an efficient, accurate, and non-invasive alternative to traditional manual scoring. By leveraging computer vision and machine learning, this approach reduces human bias, saves time, and enables consistent evaluation of body traits linked to productivity, health, and longevity. This paper presented a ResNet-50 CNN integrated with a YOLOv5 detection pipeline for cattle- buffalo image classification, achieving 93.2 % accuracy, precision 0.92, and recall 0.94 on 2,500 annotated images. The system substantially outperforms traditional machine learning approaches including SVM and KNN. Despite challenges such as variations in lighting, posture, and background, deep learning models—especially CNNs— can achieve high classification accuracy. With further development and integration into farm management systems, this technology has the potential to revolutionise livestock monitoring, improve breeding decisions, and enhance overall herd productivity.



IX. FUTURE WORK

Future research directions include:

(1) Extending the system to breed-level classification within cattle and buffalo species;

(2) Incorporating multimodal data streams such as video sequences, skeletal keypoint estimation, and weight/milk yield sensors for comprehensive evaluation

(3) Deploying the model on mobile apps or farm drones for on-field ATC; (4) expanding the dataset to cover a wider range of breeds, age groups, and geographic regions.

(4) Integrating the classification system with IoT sensor networks and cloud-based livestock management platforms for end-to-end digital agriculture.

X. REFERENCES

[1] M. E. Hossain et al., “A systematic review of machine learning techniques for cattle identification,” J. Agricultural Informatics, vol. 13, no. 2, pp. 1–16, 2022.

[2] O. Ermetin, “Deep-learning-based buffalo identification through muzzle dermatoglyphics images,” Archives Animal Breeding, vol. 68, no. 6,pp. 473–482, 2025.

[3] Taunk, D., Patel, M. (2021). Hybrid Restricted Boltzmann Algorithm for Audio Genre Classification. In: Sheth, A., Sinhal, A., Shrivastava, A., Pandey, A.K. (eds) Intelligent Systems. Algorithms for Intelligent Systems. Springer, Singapore. https://doi.org/10.1007/978-981-16-2248-9_11

[4] J. Redmon and A. Farhadi, "YOLOv3: An incremental improvement,"arXiv:1804.02767,2