

A Video Surveillance-Based Framework for Early Detection of Lumpy Skin Disease in Cattle Using Haar Cascade and Ensemble Learning

Ithaya Rani Panneer Selvam^{1*}, Akilan Selvaraj Saroja², Karuppasamy Palraj³, Parshapa Padma⁴, Sabenabanu Abdulkadhar⁵, Syed Rabiya Mohamed Ali⁶ and Hari Prasath Thinakaran⁷

¹Department of CSE, Koneru Lakshmaiah Education Foundation, Vaddeswaram, Andhra Pradesh, India

²Department Faculty of Business, University of Digital Economics and Agrotechnologies, Tashkent, Uzbekistan

³Department of Information Technology, Ramco Institute of Technology, Rajapalayam, India

⁴Department of CSE, CVR College Of Engineering VastuNagar, Mangalpalli, Ibrahimpatnam, Telengana, India

⁵Department of Computer Applications, SRM Institute of Science and Technology, Katankulathur, Chennai, India

⁶Department of CSE, Vel Tech Rangarajan Dr. Sagunthala R&D Institute of Science and Technology, Chennai, India

⁷Department of EEE, Kamaraj College of Engineering and Technology, Virudhunagar, Tamil Nadu, India

Article Information

Received 12 July 2026

Revised 25 August 2026

Accepted 10 September 2026

Available online 19 October 2026 (early access)

Authors' Contribution

IRPS: Conceptualization, ASS:

Formal analysis, KP: Software, PP:

Review, SA: Editing, SRMA: Data

collection, HPT: Methodology

Key words

Haar-like features, Ensemble method, Video surveillance, Haar cascade classifier

ABSTRACT

The growing incidence of lumpy skin disease (LSD) has increased the need for accurate and automated methods for early disease detection in cattle. This study proposes an automated video surveillance framework for continuous monitoring and early detection of LSD in cattle under real farm conditions. A fixed surveillance camera continuously monitors the cattle, and frames are extracted at regular intervals (e.g., every hour) for analysis. The extracted frames are preprocessed and augmented to improve image quality. Haar-like features are then extracted using an integral image and Haar cascade technique. Finally, an ensemble classifier distinguishes healthy cattle from LSD-infected cattle based on Haar-like features learned from the Mendeley cattle image dataset. The proposed framework achieved a classification accuracy of 95%, demonstrating its ability to provide accurate, efficient, and practical early detection of Lumpy Skin Disease in real-world farm environments.

INTRODUCTION

Livestock farming plays a vital role in global food security, yet infectious diseases such as lumpy skin disease (LSD) continue to cause significant economic losses through reduced milk yield, skin damage, and

lower animal productivity. Prompt disease identification is essential for limiting outbreaks and improving animal welfare. However, conventional diagnostic techniques often require laboratory testing, skilled personnel, and considerable time, making them less practical for continuous monitoring in large-scale or remote farming environments (Ithaya *et al.*, 2026). In contrast, AI-driven video surveillance enables continuous, non-invasive monitoring and supports timely disease detection under field conditions. Recent developments in computer vision have automated livestock health monitoring. Convolutional Neural Networks (CNNs) (Abdullah *et al.*, 2025), including lightweight architectures for real-time inference and deeper models such as DenseNet, have demonstrated strong performance in lesion recognition (Senthilkumar *et*

* Corresponding author: muhilrani@gmail.com
0030-9923/2026/0001-0001 \$ 9.00/0



Copyright 2026 by the authors. Licensee Zoological Society of Pakistan.

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

et al., 2024; Pal *et al.*, 2025). AI-based surveillance systems have been applied to cattle detection, identification, tracking, and health assessment using benchmark datasets and robust detection frameworks (Mon *et al.*, 2024; Yu *et al.*, 2024; Sharma, 2024; Araujo and Rili, 2025). Despite their effectiveness, most deep learning approaches require extensive annotated datasets and substantial computational resources, which can restrict their practical deployment in resource-constrained livestock farms (Sahare, 2024).

Handcrafted features, such as Haar-like descriptors, provided a computationally efficient alternative for real-time image analysis. By capturing local intensity variations, Haar-like features effectively represented structural characteristics and demonstrated promising performance in recent visual recognition applications (Sedaghatjoo *et al.*, 2024; ElKari *et al.*, 2024; Antipona *et al.*, 2024). Nevertheless, their application to continuous video-based detection of LSD in cattle remains limited.

To overcome this limitation, this study presents an automated video surveillance framework for early LSD detection based on the integration of Haar-like features with an AdaBoost classifier. Fixed surveillance

cameras continuously capture cattle images, from which complementary structural and texture descriptors are extracted for disease classification using Adaboost. The proposed framework provides a computationally efficient and accurate solution for real-time LSD detection, making it well suited for practical deployment in livestock farming environments.

MATERIALS AND METHODS

The overall process of the proposed video-based LSD detection system is depicted in Figure 1. The process begins with video acquisition, where five fixed high-definition surveillance cameras are strategically installed to monitor a 20 m × 30 m (600 m²) cattle shed accommodating approximately 150 cattle. The cameras continuously capture video streams, and image frames are extracted from each camera at regular intervals for analysis. For identifying LSD, the extracted video frames are compared with static cattle skin images from the trained Mendeley dataset. The dataset is subsequently partitioned into training and testing subsets to support feature

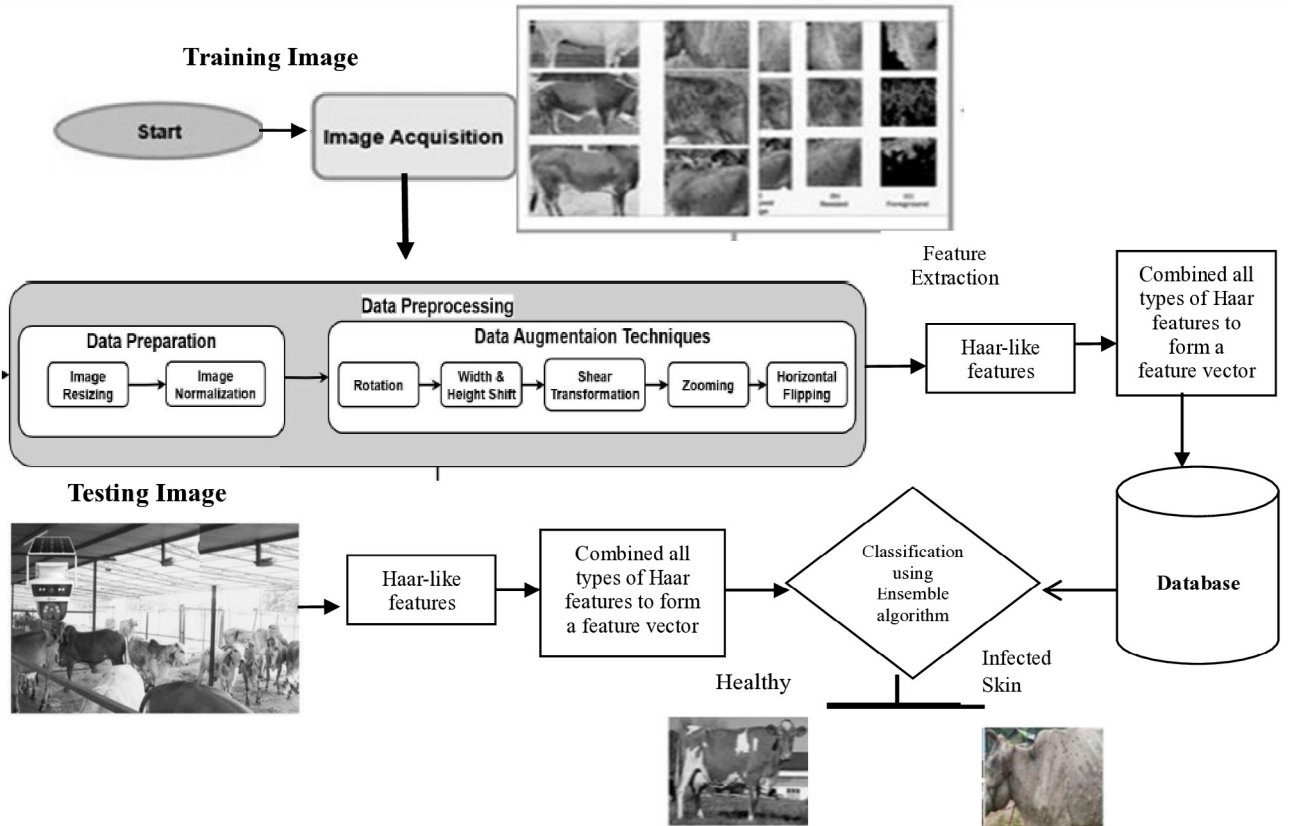


Fig. 1. Proposed video-surveillance framework for early LSD detection using Haar-like features and ensemble classification.

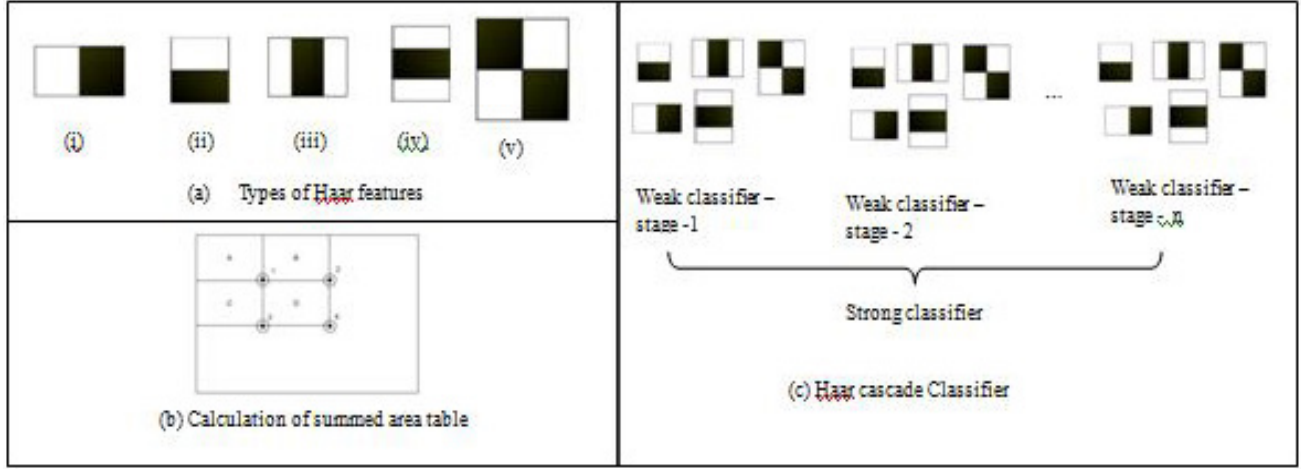


Fig. 2. Haar-like feature extraction framework, including feature types, integral-image computation and Haar cascade classification.

extraction, feature fusion, and ensemble classification. For training phase, images are first standardized via resizing and pixel normalization. To prevent overfitting and boost model robustness, data augmentation is performed using random rotations, vertical/horizontal translations, shear, zoom, and horizontal mirroring. For feature extraction and classification, Haar-like are extracted and fused into a single, cohesive feature vector which is stored in the database for model training. In the testing stage, unseen images or video frames undergo the same feature extraction protocol. These fused vectors are then classified using the AdaBoost algorithm with training database, which effectively categorizes images into healthy or infected skin states with high accuracy and low computational overhead.

Image preprocessing and data augmentation

Image preprocessing was performed to improve the reliability of LSD detection. All cattle images were resized to 240×240 pixels using bilinear interpolation, normalized to the range 0–1, and denoised to ensure consistent image quality before feature extraction. To improve model robustness and reduce overfitting, data augmentation techniques were applied, including random rotation ($\pm 25^\circ$), horizontal and vertical shifts (10%), shear transformation (0.2), random zoom (20%), and horizontal flipping (50% probability). After geometric transformations, missing boundary pixels were filled using nearest-neighbor interpolation to preserve image continuity and structural information, enabling reliable feature extraction under varying imaging conditions.

Feature extraction using Haar-like features

Haar-like features were extracted from the preprocessed cattle images to characterize intensity

variations and structural patterns associated with LSD lesions. The features measure intensity differences between adjacent rectangular regions and can capture edges, line-like structures, local contrast, and texture variations. Three configurations two-rectangle, three-rectangle, and four-rectangle Haar-like features were considered. Figure 2 illustrates the different categories of Haar-like features based on the arrangement of adjacent rectangular regions.

Categories of Haar-like features

Each image (240×240) was divided into non-overlapping (20×20) pixel windows. Within each window, five Haar-like responses were calculated: horizontal and vertical two-rectangle features, horizontal and vertical three-rectangle features, and one four-rectangle feature. The responses from all 144 windows were concatenated to form a 720-dimensional feature vector ($144 \text{ windows} \times 5 \text{ features}$). This combined representation provided complementary structural and contrast information for subsequent classification using the ensemble learning algorithms.

For computational efficiency, an integral-image representation was used to calculate rectangular intensity sums rapidly. The detailed mathematical formulations and derivations of the individual Haar-like feature types, integral-image computation, and cascade-stage classification are presented in the following sections.

Two-rectangle Haar features

A two-rectangle Haar feature is obtained by calculating the difference between the cumulative pixel intensity values of two neighboring rectangular regions with identical dimensions. These regions can be arranged either vertically or horizontally, as shown in the Figure 2a

(i). The feature value for both rectangles is calculated as:

$$Two_{Rect}Hf = \sum SP_{white} - \sum SP_{black}$$

$$Two_{Rect}Vf = \sum SP_{white} - \sum SP_{black}$$

Where $\sum SP_{white}$ and $\sum SP_{black}$

Denoted the sums of pixel intensities in the white and black rectangles, respectively. The resulting feature values for the horizontal and vertical two-rectangle configurations are represented by $Two_{Rect}Hf$ and $Two_{Rect}Vf$.

For a window size of (20×20) pixels, the vertical two-rectangle feature shown in Figure 2a(i) consists of two adjacent regions, each measuring (10×20) pixels. Likewise, the horizontal two-rectangle feature illustrated in Figure 2a(ii) is composed of two regions of (20×10) pixels each. In both cases, the feature value is obtained by subtracting the sum of pixel intensities in the black region from that of the white region. The output feature vector generated from all valid two-rectangle windows across the image is expressed as:

$$Two_{Rect}FV = [Two_{Rect}Hf_1, Two_{Rect}Hf_2, \dots, Two_{Rect}Hf_{Tot}, Two_{Rect}Vf_1, Two_{Rect}Vf_2, \dots, Two_{Rect}Vf_{Tot}]$$

Where Tot represents the total number of evaluated window.

Three-rectangle Haar features

A three-rectangle Haar feature consists of three adjacent rectangular regions arranged either vertically or horizontally, as illustrated in the Figure 2a. The feature is derived from the difference between the combined pixel intensity sums of the two outer rectangles from the pixel intensity sum of the central rectangle. Mathematically, the feature values for the horizontal and vertical three-rectangle configurations are calculated as:

$$Three_{Rect}Hf = \sum SP_{center} - (\sum SP_{left} + \sum SP_{right})$$

$$Three_{Rect}Vf = \sum SP_{center} - (\sum SP_{left} + \sum SP_{right})$$

Where $\sum SP_{center}$ denotes the sum of pixel intensities in the center rectangle, and $\sum SP_{left}$ and $\sum SP_{right}$ represent the sums in the two outer rectangles. This type of Haar feature is particularly effective for capturing line-like structures and local texture variations within an image. For a window size of (20×20) pixels, the vertical three-rectangle feature shown in Figure 2a(iii) is composed of three adjacent vertical regions, where the left and right white rectangles each measure (5×20) pixels, and the central black rectangle measures (10×20) pixels. Similarly, the horizontal three-rectangle feature illustrated in Figure 2a(iv) consists of three horizontal regions, with the top and bottom white rectangles measuring (20×5) pixels each and the middle black rectangle measuring (20×10) pixels. By evaluating these configurations over all valid window

positions, a feature vector is generated and represented as:

$$Three_{Rect}FV = [Three_{Rect}Hf_1, Three_{Rect}Hf_2, \dots, Three_{Rect}Hf_{Tot}, Three_{Rect}Vf_1, Three_{Rect}Vf_2, \dots, Three_{Rect}Vf_{Tot}]$$

Where Tot represents the total number of evaluated window.

Four-rectangle Haar features

A four-rectangle Haar feature consists of four equally sized rectangular regions arranged in a square or checkerboard pattern. The feature value is obtained by calculating the difference between the cumulative pixel intensity values of the corresponding rectangular regions. In the diagonal pairs of white and black rectangles and is expressed as:

$$Four_{Rect}F = (\sum SP_{top-left} + \sum SP_{bottom-right}) - (\sum SP_{top-right} + \sum SP_{bottom-left})$$

Where $\sum SP_{top-left}$ and $\sum SP_{bottom-right}$ denote the pixel intensity sums of the two black rectangles, and $\sum SP_{top-right}$ and $\sum SP_{bottom-left}$ (S_{b2}) represent the sums of the two white rectangles. This feature is effective for capturing local contrast variations, edge transitions, and corner-like structures within an image. As illustrated in Figure 2a(v), for a (20×20) window the four-rectangle feature divides the region into four equal (10×10) subregions arranged in an alternating checkerboard pattern of white and black rectangles. The feature response is obtained by calculating the difference between the combined pixel intensity sums of the white and black regions. By extracting these responses across all valid window positions, the resulting Haar-like descriptors effectively characterize local structural patterns, intensity contrasts, and lesion-related features, thereby providing discriminative information for automated LSD identification. The corresponding feature vector generated from all valid four-rectangle windows is represented as:

$$Four_{Rect}FV = [Four_{Rect}f_1, Four_{Rect}f_2, \dots, Four_{Rect}f_{Tot}] \dots (4)$$

where Tot represents the total number of evaluated window.

Fusion of two-, three-, and four-rectangle Haar features

To obtain a comprehensive representation of local image characteristics, two-rectangle, three-rectangle, and four-rectangle Haar-like features were extracted from each preprocessed image using a non-overlapping sliding window approach. Each input image of size (240×240) pixels was partitioned into (20×20) pixel windows, resulting in a total of 144 windows. Within every window, five Haar-like features were computed, comprising one vertical two-rectangle feature, one horizontal two-rectangle feature, one vertical three-rectangle feature, one horizontal three-rectangle feature, and one four-rectangle feature. These responses were concatenated to form a one-

dimensional feature vector. Consequently, the total feature vector size for each image was 720 features (144 windows $\times$ 5 features per window = 720 features).

$$FV = [Two_{rect} FV \cup Three_{rect} FV \cup Four_{rect} FV] \dots (5)$$

The two-rectangle features capture intensity differences between adjacent regions and are effective for detecting edges and boundaries. The three-rectangle features emphasize line-like structures and texture variations by comparing the central region with its neighboring regions. The four-rectangle features identify local contrast changes and corner-like patterns through comparisons between diagonally arranged rectangular regions. The responses from all Haar-like feature configurations were concatenated to form a unified 720-dimensional feature vector, providing complementary structural, textural, and contrast-based information. This combined feature representation enhances the discrimination of lesion-related patterns and serves as an effective input for the automated identification and classification of LSD.

Integral image generation for Haar-like feature extraction

An integral image (summed-area table) was used to efficiently calculate the intensity sums of rectangular regions during Haar-like feature extraction. This representation allows the sum of pixels within a rectangular region to be obtained from a small number of cumulative-sum values, thereby reducing the computational cost of repeated feature calculations. The detailed mathematical formulation of the integral image and rectangular-sum calculation is provided in below.

The integral image value, $II(x,y)$ at the pixel location (x,y) is expressed as:

$$II(x,y) = \sum_{x' \leq x, y' \leq y} I(x',y') \dots (6)$$

Where $I(x',y')$ denotes the pixel intensity of the original image and $II(x,y)$ represents the corresponding integral cumulative sum. This formulation enables the pixel sum of any rectangular region to be determined using only four values from the integral image, irrespective of the rectangle's dimensions. The computation procedure is illustrated in Figure 2b. Let the cumulative sums at locations 1, 2, 3, and 4 correspond to regions A, A+B, A+C, and A+B+C+D, respectively. The sum of pixel intensities within rectangle (D) can then be obtained efficiently as $4 + 1 - (2 + 3)$. By eliminating the need to repeatedly sum individual pixels within each region, the integral image substantially reduces the computational complexity of rectangular feature evaluation.

The computed integral image was subsequently employed to calculate the rectangular sums required

for two-rectangle, three-rectangle, and four-rectangle Haar-like features. This efficient computation strategy accelerates feature extraction while preserving accuracy, making it well suited for automated LSD identification from cattle images.

Haar cascade classifier for LSD detection

Following the extraction of Haar-like features, the Haar cascade classifier was utilized to identify image regions exhibiting characteristics associated with LSD. The Haar cascade framework employs a sequence of trained classifiers arranged in multiple stages, where each stage evaluates the extracted Haar-like feature responses to distinguish relevant patterns from non-relevant regions. This cascaded structure enables rapid rejection of regions that do not contain disease-related features while focusing computational resources on potentially affected areas. By combining multiple Haar-like features and decision rules, the Haar cascade classifier effectively detects characteristic lesion patterns, nodules, and texture abnormalities associated with LSD, thereby supporting accurate and computationally efficient disease identification as shown in Figure 2(c).

$$H(x,y) = FV \leq T \dots (7)$$

Here, T is the threshold for classification.

By combining multiple Haar-like features with the AdaBoost classifier, the Haar cascade algorithm effectively learns discriminative patterns for accurate object recognition. The training process involves a set of weak classifiers $h_j(x)$, where (x) denotes the extracted Haar-like feature vector. Each weak classifier evaluates a single feature and classifies the input as either LSD-positive (+1) or healthy (-1) based on a predefined threshold. The most discriminative features are selected and combined through the AdaBoost algorithm to construct a strong classifier. These strong classifiers are then arranged in a cascade structure, enabling rapid elimination of non-relevant regions while focusing on potentially affected areas. This hierarchical approach improves computational efficiency and supports accurate automated identification of lumpy skin disease.

$$h_j(x) = \begin{cases} +1 & \text{if } p_j(x) \leq \theta_j \quad \text{LSD} \\ -1 & \text{healthy} \end{cases} \dots (8)$$

Where $p_j(x)$ denotes the output of the j -th weak classifier for the input sample x and θ_j represents its associated threshold. The final strong classifier, $H(x)$ is constructed by linearly combining the outputs of the selected weak classifiers according to their respective weights:

$$H(x) = \sum_{j=1}^T \alpha_j h_j(x) \dots (9)$$

Where T denotes the total number of weak classifiers, and α_j denotes the weight assigned to the j -th weak classifier according to its classification performance. During the boosting procedure, the weights of the training samples are iteratively updated to emphasize instances that are difficult to classify.

$$w_i^{(t+1)} = w_i^t \cdot \exp(-\alpha_t y_i h_t(x_i)) \dots (10)$$

Where y_i represents the ground-truth label of the i -th training sample, and t denotes the current boosting iteration. The AdaBoost algorithm assigns an importance weight to each weak classifier, as defined in below equation.

$$\alpha_t = \frac{1}{2} \ln \frac{1 - \epsilon_t}{\epsilon_t}$$

where ϵ_t denotes the weighted classification error of the j -th weak classifier. The trained cascade classifier, organized into (S) sequential stages, is applied using a sliding-window approach during the detection phase. The window systematically traverses the input image at multiple locations to identify regions exhibiting characteristics of LSD.

$$H(x) = \sum_{s=1}^S \sum_{j=1}^{T_s} \alpha_j^{(s)} h_j^{(s)}(x) \dots (11)$$

Where S represents the total number of cascade stages, and T denotes the number of weak classifiers included in the S -th stage. The strong classifiers are organized sequentially in a cascade structure, allowing each stage to evaluate only those image regions that successfully pass through all preceding stages.

Ensemble-based classification framework

Ensemble learning combines the predictions of multiple models to improve classification accuracy and robustness. The boosting framework by [Khotimah \(2026\)](#) is selected for its efficacy in minimizing false-positive rates. Beyond improving accuracy, it inherently performs feature selection by identifying the most discriminative features within the high-dimensional space, reducing redundancy and computational complexity. However, standard boosting remains susceptible to performance degradation from noisy samples and extreme outliers.

Boosting variants for disease recognition

To systematically evaluate the impact of different boosting paradigms, this study implements three prominent variants:

(i) Real AdaBoost (RA) - replaces rigid binary classifications with continuous, confidence-based probability estimates. While it rapidly minimizes exponential loss to extract powerful features, its aggressive optimization makes it vulnerable to dataset noise.

(ii) Gentle AdaBoost (GA) - uses a conservative, Newton-step update mechanism that bounds the regression outputs of the weak learners. This smoother approach shields the model from being skewed by outliers, skin artifacts, or heavy shadows.

(iii) Modest AdaBoost (MA) ([Vezhnevets et al., 2005](#)) - features an advanced regularization technique that grades weak learners against the inverse distribution of the data. By downscaling learners that struggle with previously settled samples, it prevents overfitting and yields the highest reliability on unseen field data.

These ensemble architectures are trained utilizing the combined Haar-ILBP feature vectors to perform binary classification, distinguishing between healthy and LSD-infected cattle skin images.

RESULTS AND DISCUSSION

To validate the video surveillance-based LSD detection system under practical farm conditions, five fixed high-definition surveillance cameras were installed at Yash Farms, Gunadala, Vijayawada, covering a $(20\text{m} \times 30\text{m})$ (600 m^2) cattle shed with approximately 150 cattle (including 138 healthy and 12 LSD-infected). Image frames extracted at regular intervals from each camera and were used to evaluate the proposed model. The extracted frames were resized to (240×240) pixels and included different cattle breeds, including Jersey and indigenous (Desi) cattle, with variations in age and field conditions. Knowledge learned from these datasets was applied to classify the surveillance image frames.

The publicly available datasets were used for training and evaluation such as: the Mendeley LSD dataset ([Kumar and Shastri, 2022](#)). The data set contained 1023 images, including 478 healthy and 545 LSD-infected cattle images with a resolution of 256×256 pixels. It included multiple cattle breeds, such as Santa Gertrudis, Brahman, Guzerat, and Yak, captured under varying poses, illumination conditions, and viewing perspectives. The proposed framework was implemented using MATLAB R2017 on an Intel Core i5 (3.20 GHz) processor with 8 GB RAM. Although this demonstrates its suitability for modest computational environments, the software and hardware configuration may limit reproducibility and scalability. Future work will evaluate the framework using modern platforms such as PyTorch and TensorFlow. The Mendeley dataset were divided into 80% for training and 20% for testing.

Performance analysis of Haar-like features

The classification performance was evaluated using different Haar-like features with different adaboost techniques. Two-rectangle features captured edge information, three-rectangle features represented line and

Table I. Classification performance of individual and combined Haar-like feature types using Real Adaboost (RA), Gentle Adaboost (GA), and Modest AdaBoost (MA) classifiers.

Types of Haar- features	Feature vector size	Accuracy (%)					
		Mendeley dataset			Yash farm surveillance data		
		RA	GA	MA	RA	GA	MA
Two-rectangle features (both horizontal + vertical)	288	89	90.4	90	95	97.4	96
Three-rectangle features (both horizontal + vertical)	288	90	91	90.4	96	98	99
Four-rectangle features	144	92	91	93	97	98	97
Combined features	720	93	95.11	93	97	99	97

intensity transitions, and four-rectangle features extracted complementary texture characteristics. Combining all feature types produced a 720-dimensional feature vector and achieved the highest classification accuracy (Table I) with GA, demonstrating that integrating complementary edge, line, and texture information provides a more discriminative representation of LSD lesions.

Performance analysis of different window size

The performance of the proposed method was evaluated using 10×10 , 20×20 , and 30×30 non-overlapping windows. The 10×10 window produced a large number of features, increasing computational complexity, while the 30×30 window failed to capture important local lesion details. The 20×20 window provided a balance between feature representation and computational efficiency with GA classifier, resulting in the highest classification performance in the datasets as shown in Table II. Therefore, the 20×20 window size was selected for subsequent experiments.

Table II. Effect of different non-overlapping Haar feature extraction window sizes.

Different window size (Pixels)	Feature vector size	Accuracy (%)					
		Mendeley dataset			Yash farm surveillance data		
		RA	GA	MA	RA	GA	MA
10×10	2880	92	94	93	95	98	96
20×20	720	93	95.1	94	97.6	99	98
30×30	320	92	93	94	97	98	97

Performance analysis of the Haar cascade classifier across different cascade stages

As the number of Haar cascade stages increased, the classification performance improved consistently, with accuracy rising from 90.80% in Stage 1 (50 weak classifiers) to 95.11% in Stage 5 (450 cumulative weak classifiers among 720 Haar-like features), accompanied by corresponding improvements in precision, recall, and

F1-score as shown in (Table III). The final model correctly classified 523 of 545 infected images (22 false negatives) and 450 of 478 healthy images (28 false positives), achieving an accuracy of 95.11%, precision of 94.92%, recall of 95.96%, and F1-score of 95.44%. These results demonstrate the robustness of the proposed approach on a moderately imbalanced dataset. The confusion matrix is shown in Table IV.

Table III. Classification performance at different Haar cascade stages showing the number of weak classifiers, cumulative classifiers, accuracy, precision, recall, and F1-score on the Mendeley LSD dataset.

Cascade stage	Weak classifiers in stage	Cumulative weak classifiers	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)
Stage 1	50	50	90.81	93.28	89.17	91.18
Stage 2	50	100	92.08	93.77	91.19	92.47
Stage 3	100	200	93.74	94.62	93.58	94.10
Stage 4	100	300	94.62	95.04	94.86	94.95
Stage 5	150	450	95.11	94.92	95.96	95.44
Stage 6	100	550	94.04	94.49	94.31	94.40
Stage 7	100	650	92.67	93.84	92.29	93.06

Table IV. Confusion matrix showing the true-positive, false-negative, true-negative, and false-positive classifications obtained at each Haar cascade stage on the Mendeley LSD dataset.

Stage	True positive	False negative	True negative	False positive
1	486	59	443	35
2	497	48	445	33
3	510	35	449	29
4	517	28	451	27
5	523	22	450	28
6	514	31	448	30
7	503	42	445	33

Table V. Performance of the proposed Haar-like features and different AdaBoost classifier - based video surveillance framework for LSD detection using real-farm video data from Yash Farm.

Testing through video camera	Model name	Accuracy (%)	F1 score	Recall	Precision	Training dataset
Yash Farm	Haar like features with RA	96	95.5	96	95	Mendeley dataset
Yash Farm	Haar like features with MA	97	96.5	97	96	Mendeley dataset
Yash Farm	Haar like features with GA	99	97.7	98	97.5	Mendeley dataset

Performance evaluation using real-farm video validation

The proposed framework was validated using video data collected from Yash Farm. Three classifiers were evaluated using combined rectangle Haar-like features and models trained on the Mendeley dataset. RA, MA, and GA AdaBoost achieved accuracies of 96%, 97%, and 99%, respectively as shown in Table V. Among them, Gentle AdaBoost provided the highest performance, with an F1-score of 97.7%, 98% recall, and 97.5% precision, demonstrating its suitability for real-time LSD detection under practical farm conditions.

Performance comparison of state-of-the-art models for LSD detection

To evaluate the relative effectiveness of the proposed framework, a direct experimental comparison was performed using the same Mendeley LSD dataset and an identical 80:20 training–testing protocol. The proposed Haar-like feature and Gentle AdaBoost framework was compared with ResNet, VGG16 + ANN, CNN + InceptionV3, CNN, Xception, DenseNet121, Vision Transformer, YOLO + EfficientNet, and VGG16 + ResNet. As reported in Table VI, the proposed approach achieved the highest accuracy of 95.11% with an execution time of approximately 22 s, while the evaluated deep learning and hybrid approaches achieved accuracies between 89.0% and 94.2%. The 10-fold cross-validation accuracy of $96.2 \pm 0.8\%$ further indicates the stability of the proposed framework.

The proposed model was evaluated using an 80:20 training–testing split and further validated through statistical analyses. It achieved a Cohen's kappa coefficient of 0.96 (95% CI: 0.93–0.96), indicating almost perfect agreement with the ground truth. McNemar's test showed no significant difference in paired misclassification patterns compared with the baseline ($p > 0.05$), while a paired t-test confirmed that the proposed Haar like features fusion with AdaBoost significantly outperformed the baseline methods ($p < 0.05$). Furthermore, one-way ANOVA revealed significant differences among the evaluated feature extraction techniques ($p < 0.05$), demonstrating the robustness and effectiveness of the proposed framework.

Table VI. Comparative performance of the proposed Haar-like feature with Gentle AdaBoost framework and existing LSD detection approaches evaluated on the same Mendeley dataset, including classification accuracy and execution time.

Authors/ Year	Model/ Technique	Mendeley data set	
		Accu- racy	Sec- onds
Goceri <i>et al.</i> , 2019	ResNet	89	18
Simonyan <i>et al.</i> , 2015	VGG16 + ANN	90.5	22
Szegedy <i>et al.</i> , 2024	CNN + InceptionV3	91	27
Sahare, 2024	CNN	89.4	21
Chollet <i>et al.</i> , 2017	Xception with CNN	91.8	23
Abdullah <i>et al.</i> , 2025	DenseNet121	90.6	22
Ullah <i>et al.</i> , 2025	Vision Transformer (ViT)	94.2	22
Ubaidullah <i>et al.</i> , 2026	YOLO + EfficientNet	92	25
Sentoso <i>et al.</i> , 2025	VGG16 + ResNet	93.2	25
Tan and Le, 2021	EfficientNetV2	94.3	26
Liu <i>et al.</i> , 2022	ConvNeXt	95	24
Our proposed, 2026	Haar features gentle adaboost	95.11	22

These results indicate that the observed improvement was statistically meaningful; however, statistical significance should be interpreted together with the classification performance and practical relevance of the proposed framework. In particular, the real-farm validation involved approximately 150 cattle, including 138 healthy and 12 LSD-infected cattle, and the relatively small number of infected cases limits the generalizability of the findings. Therefore, larger and more diverse multi-farm datasets are required in future studies to obtain more precise estimates, stronger effect-size assessments, and broader validation of the proposed system.

Economic feasibility and veterinary integration

The proposed framework has potential for cost-effective deployment because it combines fixed surveillance cameras with lightweight Haar-like features

and Gentle AdaBoost, requiring relatively modest computational resources. However, deployment costs may include cameras, computing devices, connectivity, installation, maintenance, and technical support, while farmer adoption may depend on system reliability, awareness, and technical expertise. The system could also be integrated with veterinary extension services to generate alerts containing suspected LSD images, camera identification, and timestamps for veterinary verification and timely intervention. Future work will focus on field-based economic assessment, farmer acceptance, and integration with digital veterinary health platforms.

Scalability and practical deployment

Although continuous video surveillance can generate substantial amounts of data, the proposed framework reduces this burden by extracting representative frames at predefined intervals rather than processing and storing every video frame. This sampling strategy helps reduce storage and computational requirements while retaining relevant visual information for disease monitoring. In addition, the use of Haar-like features with Gentle AdaBoost provides a relatively lightweight classification pipeline. For larger-scale deployment, the system can be further optimized through adaptive frame-extraction intervals, edge-based processing, distributed computation, and selective retention of relevant frames or extracted feature vectors. These strategies can support deployment across multiple cameras and larger cattle populations while limiting increases in computational and storage requirements.

Deployment considerations and pilot validation

The practical feasibility of the proposed framework was evaluated through a pilot deployment at Yash Farm, where five fixed cameras were used to monitor approximately 150 cattle within a 20 m × 30 m cattle shed. The pilot demonstrated the feasibility of continuous video-based monitoring under real farm conditions and provided preliminary evidence of the system's applicability to routine cattle surveillance. From a cost-benefit perspective, the potential advantages include reduced dependence on continuous manual inspection, earlier identification of suspected LSD cases, and improved support for timely veterinary intervention. However, the overall economic benefit will depend on factors such as camera and computing costs, installation and maintenance requirements, electricity consumption, labour costs, and veterinary expenses. A comprehensive multi-farm cost-benefit analysis is therefore required to quantify the economic benefits and assess the long-term viability of the system before large-scale commercialization.

Dataset limitations and practical applicability

The current study was evaluated using the available Mendeley cattle image dataset together with real-farm surveillance data. Although the results demonstrate the feasibility of the proposed approach, we acknowledge that extreme illumination, severe occlusion, substantial background variation, significant animal movement, and different camera angles were not exhaustively evaluated. Broader validation using data collected across multiple geographical locations, cattle breeds, and seasons is therefore required to establish the generalizability of the proposed framework. Future work will focus on developing a larger, geographically diverse dataset with expert-validated annotations covering different breeds, seasonal conditions, lighting variations, occlusion levels, and farm environments. Evaluation using larger and more diverse real-farm video datasets under challenging environmental conditions will further support rigorous cross-location, cross-breed, and cross-environment validation and improve the robustness and practical applicability of the proposed video-based LSD detection system.

The proposed video surveillance framework offers a cost-effective approach for continuous cattle monitoring using fixed surveillance cameras and modest computational resources. Its automated detection capability can reduce the need for continuous manual observation and support early identification of suspected LSD cases. The real-farm validation at Yash Farms, Gunadala, Vijayawada, demonstrates the practical feasibility of deploying the system under real-world cattle-monitoring conditions. Future work will further evaluate the cost-benefit aspects and deployment across multiple farms.

CONCLUSION

The proposed video surveillance framework enables continuous and automated monitoring for early detection of lumpy skin disease (LSD) under real farm conditions. Fixed surveillance cameras captured cattle images, which were analyzed using the proposed Harr feature with Gentle AdaBoost framework. Image preprocessing and augmentation improved robustness, while Haar-like features captured lesion structures and local texture information. The framework was validated using video frames captured at Yash Farm with trained the Mendeley dataset. It achieved accuracies of 99%, demonstrating robust performance under practical farm conditions. The proposed system supports non-invasive, real-time disease monitoring, enabling early identification of infected cattle, timely intervention, reduced disease transmission, and improved herd management.

Future work will focus on integrating deep learning

and vision transformer models, expanding real-world video datasets, and developing lightweight edge-based implementations for practical on-farm deployment.

DECLARATIONS

Acknowledgments

We acknowledge the support of Koneru Lakshmaiah Education Foundation for providing resources for this research.

Funding

This research received no specific grant.

Ethical statement

This study involved non-invasive video surveillance of cattle under real farm conditions. No animals were handled, restrained, treated, or subjected to invasive experimental procedures. The study also used publicly available cattle image data for model training and evaluation.

Data availability

All data were provided in the manuscript.

Generative AI and AI assisted technology statement

The authors declare that no generative AI and AI assisted technology was used in the creation of this manuscript.

Statement of conflict of interest

The authors have declared no conflict of interest.

REFERENCES

- Abdullah, W., Tanwar, S. and Abouhawwash, M., 2025. Deep learning based detection of lumpy skin disease in livestock using CNNs. *Sustain. Mach. Intellig. J.*, **11**: 1-10. <https://doi.org/10.61356/SMIJ.2025.11515>
- Antipona, C.A., Magsino, R.R. and Dioses, R.M., 2024. An enhancement of Haar cascade algorithm applied to face recognition for gate pass security. *Proc. Int. Conf. Inf. Technol. Secur., (IC-ITECHS)*. <https://doi.org/10.2139/ssrn.5009574>
- Araujo, V.M. and Rili, I., 2025. AI-powered cow detection in complex farm environments. *arXiv Preprint*, **arXiv**:2501.02080. <https://doi.org/10.1016/j.atech.2025.100770>
- Chollet, F., 2017. Xception: Deep learning with depthwise separable convolutions. *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, pp. 1251-1258. <https://doi.org/10.1109/CVPR.2017.195>
- Elkari, B., Ourabah, L., Sekkat, H. and Farah, G., 2024. Implementation of presence detection with Haar cascade and local binary patterns histograms. *Mathemat. Model. Comp.*, **11**: 1093-1105. <https://doi.org/10.23939/mmc2024.04.1093>
- Goceri, E., 2019. Deep learning-based classification of skin lesions using ResNet architectures. *Biomed. Signal Process. Contr.*, **51**: 1-8.
- Ithaya, R.P. and Tabbassum, S., 2026. An enhanced deep learning approach for the automated identification of lumpy skin disease in cattle. *Open Vet. J.*, **16**: 3717-3730. <https://doi.org/10.5455/OVJ.2026.v16.i6.44>
- Khotimah, B.K., 2026. A robust ensemble deep learning model for lumpy skin disease identification. *Engineer. Technol. appl. Sci. Res.*, **16**: 31285–31293. <https://doi.org/10.48084/etasr.14823>
- Kumar, S. and Shastri, S., 2022. Lumpy skin images dataset. Mendeley Data 2022, V1. <https://data.mendeley.com/datasets/w36hpf86j2/1>
- Liu, Z., Mao, H., Wu, C.Y., Feichtenhofer. 2022. *A ConvNet for the 2020s*. Proceedings of the IEEE/CVF conference on computer vision and pattern recognition (CVPR), pp. 11976–11986. <https://doi.org/10.1109/CVPR52688.2022.01167>
- Mon, S.L. and Onizuka, T., 2024. AI-enhanced real-time cattle identification system through tracking across various environments. *Sci. Rep.*, **14**: 17779. <https://doi.org/10.1038/s41598-024-68418-3>
- Pal, M., Behera, S., Mohapatra, R., 2025. Implementation of a deep learning system for detection and classification of lumpy skin disease in cattle. *Vet. Med. Sci.*, **11**: e70664. <https://doi.org/10.1002/vms3.70664>
- Sahare, D.K., 2024. Extensive investigation of convolutional neural network architectures for the diagnosis of Lumpy Skin Disease in dairy cattle. *Halcyon*, **10**: e27893. <https://doi.org/10.1016/j.heliyon.2024.e34242>
- Sedaghatjoo, Z., Hosseinzadeh, H. and Shirzadi, A., 2024. Enhanced facial feature extraction and recognition using optimal fully dispersed Haar-like filters. *arXiv Preprint*, **arXiv**: 2404.10476.
- Senthilkumar, C., Sindhu, C., Vadivu, G. and Neethirajan, S., 2024. Early detection of lumpy skin disease in cattle using deep learning: A comparative analysis of pretrained models. *Vet. Sci.*, **11**: 510. <https://doi.org/10.3390/vetsci11100510>
- Sentoso, F., Ardiansyah, R. and Wangsa, S., 2025. Performance evaluation of VGG and ResNet architectures for cattle disease image classification.

- Sistemasi: J. Sistem Inf.*, **14**: 89-98.
- Sharma, A., 2024. Universal bovine identification via depth data and deep metric learning. *arXiv preprint*, **arXiv**: 2404.00172.
- Simonyan, K. and Zisserman, A., 2015. Very deep convolutional networks for large-scale image recognition. *Proc. Int. Conf. Learn. Represent. (ICLR)*.
- Tan, M. and Le, Q.V., 2021. EfficientNetV2: Smaller models and faster training. *Proc. 38th Int. Conf. Mach. Learn.*, **139**: 10096–10106.
- Ubaidullah, A., Hussain, H. and Jafri, M., 2026. LUMPNet: A hybrid YOLO-EfficientNet framework for lumpy skin disease detection. *Comput. Biol. Med.*, **180**: 107694.
- Ullah, A., 2025. Vision transformer based automated classification of cattle skin diseases. *Expert. Syst. Appl.*, **235**.
- Vezhnevets, A. and Vezhnevets, V., 2005. *Modest AdaBoost-teaching AdaBoost to generalize better*, *Graphicon-Novosibirsk Akademgorodok*.
- Yu, P. and Burghardt, T., 2024. MultiCamCows2024: A multi-view image dataset for AI-driven Holstein-Friesian cattle re-identification on a working farm. *arXiv preprint*, **arXiv**: 2410.12695.

Online First Article