A Robust Ensemble Deep Learning Model for Lumpy Skin Disease Identification

Bain Khusnul Khotimah
Department of Informatics Engineering, University of Trunojoyo Madura, Bangkalan, Indonesia
bain@trunojoyo.ac.id (corresponding author)

Budi Dwi Satoto
Department of Informatics Engineering, University of Trunojoyo Madura, Bangkalan, Indonesia
budids@trunojoyo.ac.id

Yoga Dwitya Pramudita
Department of Informatics Engineering, University of Trunojoyo Madura, Bangkalan, Indonesia
yoga@trunojoyo.ac.id

Deshinta Arrova Dewi
Center for Data Science and Sustainable Technologies, INTI International University, Malaysia
deshinta.ad@newinti.edu.my

Firdatul A'yuni
Department of Informatics Engineering, University of Trunojoyo Madura, Bangkalan, Indonesia
firdatulayuni@gmail.com

Received: 16 September 2025 | Revised: 27 October 2025 | Accepted: 12 November 2025

 DOI: https://doi.org/10.48084/etasr.14823

ABSTRACT

Early diagnosis of Lumpy Skin Disease (LSD) in cattle is crucial for maintaining meat quality and ensuring livestock productivity. This study introduces an innovative Ensemble Deep Learning (EDL) framework with Contrast Limited Adaptive Histogram Equalization (CLAHE) preprocessing, L2 regularization-based feature optimization, and adaptive ensemble weighting to efficiently and accurately detect LSD. Comparative experiments were conducted with feature selection methods, such as Information Gain (IG), Chi-square, and L1 regularization, as well as ensemble methods, such as Simple Averaging, Majority Voting, and Stacking. The L2-norm-based proposed ensemble attained the highest Dice score of 97.50%, the highest Jaccard index of 98.20%, and the best AUC of 0.980 at $\alpha = 0.01$. It also attained an accuracy of 99.70%, precision of 98.73%, recall of 96.20%, and a 97.80% F1-score, with a moderate Inference Time (IT) of 0.17 s per image. In comparison to marginally slower but simpler ensemble methods, it also exhibited large improvements in robustness, generalization, and feature stability. Overall, when combined, the ensemble optimization and feature regularization enable the Regularized-ensemble Convolutional Neural Network (RoFR-eCNN) to deliver precise, accurate LSD detection, providing a scalable and efficient framework for automatic and early livestock disease diagnosis.

Keywords-LSD; CLAHE; L2 regularization; CNN architectures; ensemble deep learning; feature selection; bioinformatics; food security; weight ensemble

I. INTRODUCTION

LSD is harmful to bovine health and welfare and causes severe economic losses in the livestock industry [1]. LSD is transmitted through both hematophagous insect vectors and direct contact with contaminated secretions and excretions from infected livestock [2]. Identification of the disease in previously disease-free regions, particularly in a herd of yaks, requires urgent surveillance and imposition of control programs to limit its spread. Climate change poses severe risks to animal health by altering ecosystems and weather patterns, and making their survival more difficult [3]. While this new threat calls for an immediate need to establish effective and rapid diagnostic methods for early-stage detection, it aligns with Sustainable Development Goal (SDG) 2: Zero Hunger by safeguarding livestock health and food security [4]. Early identification of

the disease is key to limiting its spread and optimizing the effectiveness of control strategies, thus ensuring livestock health and food quality and security.

According to [5], the application of an adjusted Capsule Network (CapsNet) is essential for early LSD detection using Artificial Intelligence (AI). A Machine Learning (ML) self-learning algorithm was designed in an attempt to detect LSD using Inception V3 as a cattle image feature extractor. These characteristics were subsequently classified using Support Vector Machines (SVM) [6]. However, Deep Learning (DL) models, developed using pre-trained networks, have been extremely accurate in classification, and they exhibit strong potential as fast, non-invasive diagnostic techniques to facilitate efficient field-level disease management [7].

Approaches to address the challenge of developing an EDL model for LSD detection using feature selection and algorithm optimization have been explored. A Convolutional Neural Network (CNN), which included dataset collection, preprocessing, and parameter tuning, was developed for the LSD detection in cattle [8], revealing reliable model performance. The study presents an automatic cattle LSD detection system based on DL with several pre-trained CNN models, including VGG16 and MobileNetV2. The models were compared based on transfer learning along with data preprocessing, augmentation, and class balancing, all of which helped in achieving accuracy and improving the generalization capacity of the model [9]. A comparative analysis of two popular CNN models, Inception and Xception, for the classification of images of LSD infected cattle showed that Xception outperformed Inception with an accuracy of 98.8% [10]. However, MobileNetV2 combined with the RMSProp optimization algorithm [11], and transfer learning models ResNet-50 and VGG-16 have been some of the most successful architectures for LSD classification. Meanwhile, the combination of the Convolutional Block Attention Module (CBAM) [12] with DenseNet has been shown to enhance the ability of the network to emphasize salient features, thereby improving the effectiveness of attention-based CNN approaches in animal disease detection.

Another optimization method for LSD classification used feature selection with FeatureWiz for ML algorithms and compared two main methods, the stacked ensemble and the optimized Artificial Neural Network (ANN) [13]. The results indicated that the optimized ANN achieved slightly better results compared to the stacked ensemble. Feature selection with dimensionality reduction through PCA [14] was performed by combining skin area segmentation through color histogram and feature extraction through a 10-layer CNN. The method achieved an accuracy of up to 96%, highlighting its superior effectiveness compared to conventional approaches. Furthermore, the combination of feature selection techniques that are based on high correlation [15] with ensemble learning algorithms can significantly enhance the performance of ML models in detecting LSD in cattle. Improvements in the preprocessing stages to improve the quality of the images [16] and in highlighting the area of interest have effectively led to developing a model that is able to classify the early stages of LSD in cattle, where the ensemble method used has led to high accuracy.

To address these shortcomings, the current study introduces a new approach that integrates feature selection into an ensemble CNN framework by exploiting majority voting [17]. In addition, CLAHE preprocessing, which was initially applied on pet skin disease detection, has proven effective in enhancing lesion contrast and is utilized in this study to enhance LSD image classification [18].

The proposed RoFR-eCNN framework integrates four pre-trained CNN architectures, namely DenseNet121, ResNet50, InceptionV3, and MobileNetV2, through transfer learning to leverage complementary feature representations. Both DenseNet121 and InceptionV3 have generic low-level features, ResNet50 reinforces global and local context learning by bottleneck designs, and MobileNetV2 offers computational efficiency by depth-wise separable convolutions. In contrast to previous approaches, where CNN feature extraction and preprocessing were treated as separate stages, the proposed RoFR-eCNN integrates CLAHE with L2-regularized convolutional layers (λ) for effective feature extraction and implicit suppression of redundant activations, resulting in enhanced discriminative power for LSD image classification. Besides, an L2-constrained ensemble weight optimization (α) is introduced to balance the contributions of heterogeneous CNN models for generalized and stable ensemble predictions beyond conventional majority voting. The proposed framework, thus, devises a lightweight and yet strongly discriminative diagnostic model with robustness to diverse image conditions without affecting computational efficiency. This study offers the following key contributions:

- Enhanced feature extraction and selection by integrating CLAHE preprocessing with L2-regularized CNN layers (λ) for a more discriminative representation for LSD identification.
- A weighted ensemble CNN framework integrating various pre-trained models with L2-constrained weight optimization (α) to achieve robust and well-balanced model fusion.
- Deployment of depthwise separable convolutions and hybrid architectural designs for achieving high efficiency while maintaining top-class classification performance.

II. METHODOLOGY

This study presents a robust segmentation-based approach to enhanced classification accuracy in early LSD detection in cattle. The methodology begins with preprocessing, which is an end-to-end pipeline consisting of a series of definitive stages: resizing, CLAHE, augmentation, and normalization, for ensuring uniform inputs. Preprocessing begins with CLAHE tailored to enhance image contrast and highlight local features. Feature extraction is performed using CNN architectures with L2 regularization to reduce feature redundancy and emphasize discriminative features. The extracted features are subsequently integrated into an EDL framework using a weighted averaging strategy to improve stability and generalization. The model performance is validated using accuracy, precision, recall, F1-

score, specificity, Jaccard index, and Dice coefficient. Figure 1 illustrates the overall process of the proposed RoFR-eCNN method for LSD detection.

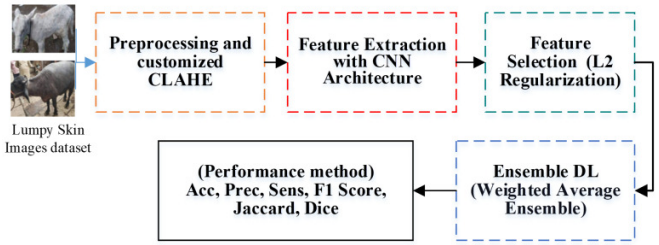

Fig. 1. Process of the proposed RoFR-eCNN method for LSD detection.

A. Dataset Collection

The LSD image dataset consists of 1,024 image records categorized into two classes, namely Lumpy Skin (324 images) and Normal Skin (700 images) [19]. All images were preprocessed in the same PNG format with a common resolution of 256×256 pixels. The Lumpy Skin class depicts cattle infected by LSD, demonstrating common skin nodules with a diameter of 2–5 cm, swelling (edema), and, in some measure, ulceration. In contrast, the Normal Skin class contains images of healthy cows with no symptoms of the disease. The dataset supports computer vision–based image classification research, particularly for the automatic detection and diagnosis of LSD in cows. Figure 2 displays sample images from the classes.

Fig. 2. Sample images of the dataset: (a) Normal Skin class, (b) Lumpy Skin class.

B. Data Preprocessing

The preprocessing stage began with resizing from the original 256×256 pixels to 224×224 pixels to render image sizes consistent. Subsequently, CLAHE was applied, which splits the image into small blocks and performs histogram equalization on each block individually [20].

The data augmentation step was also incorporated for diversification training on samples with three methods used: rotation at random to change image orientation, brightness to alter light intensity levels, and zoom-in to focus on particular areas in an image. Min–Max normalization scales pixel intensity values into the interval [0, 1] without loss of contrast, enabling faster and more stable convergence for DL models. Table I presents the data size before and after augmentation.

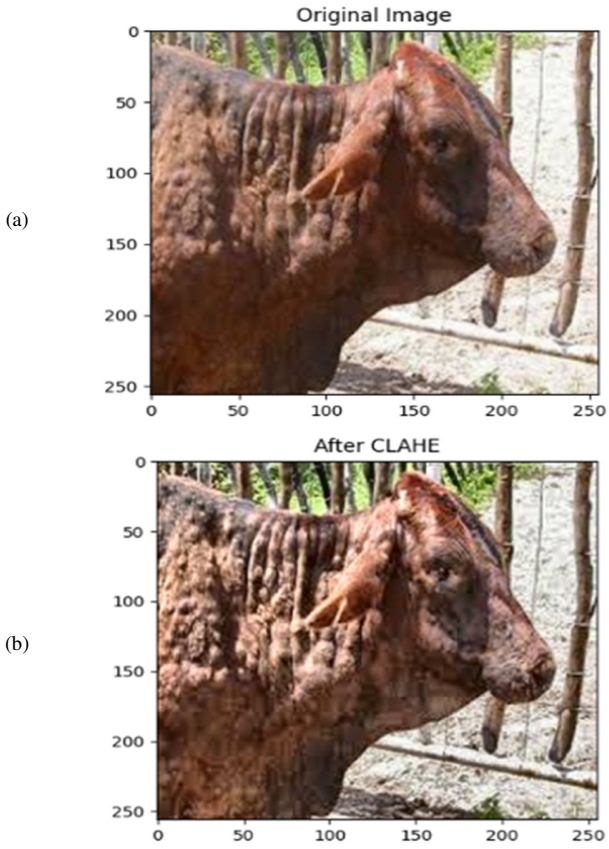

Fig. 3. Data preprocessing with CLAHE: (a) original image, (b) image after CLAHE preprocessing.

TABLE I. TOTAL SIZE OF THE DATASET

Category	Before augmentation	After augmentation
Normal Skin	700	777
Lumpy Skin	324	1704
Total	1024	2481

TABLE II. DATASET SIZE AFTER AUGMENTATION

Class	Training	Validation	Testing
Normal Skin	777	71	71
Lumpy Skin	1704	32	33

C. Data Split

The dataset was split into three distinct subsets: training, validation, and testing, using an 80:10:10 ratio. The data were obtained after the augmentation process to observe the distribution based on position. Table II depicts the dataset distribution used to evaluate the method's performance, with 80% allocated for training, 10% for validation, and 10% for testing.

D. Contrast Limited Adaptive Histogram Equalization

CLAHE enhances local contrast by dividing an image into very small patches and equating each area independently without amplifying noise in homogeneous regions. Although no prior study has specifically applied CLAHE directly to LSD image datasets, its effectiveness has been demonstrated for

other similar applications. For instance, authors in [18] utilized CLAHE to enhance lesion contrast for the detection of pet skin disease, while authors in [20] reported its efficiency in improving image quality for the classification of diabetic retinopathy. This evidence supports the potential of CLAHE to enhance contrast in cattle skin images for LSD detection. Unlike global histogram equalization, CLAHE prevents noise in uniform areas by applying a clip limit to the histogram to achieve fair brightness and contrast. In this approach, the intensity value $I'_{(x,y)}$ of each pixel is computed using:

$$CDF_k = \sum_{i=0}^{k} H_i^{clip} \quad (1)$$

$$I'_{(x,y)} = (L-1) \times \frac{CDF(I(x,y)-CDF_{min})}{MN-CDF_{min}} \quad (2)$$

where $I'_{(x,y)}$ represents the new intensity of the pixel at coordinates (x, y), L denotes the number of gray levels in the image, $CDFI'_{(x,y)}$ is the cumulative distribution function value of the pixel intensity, CDF_{min} is the minimum CDF value used to normalize the scale, and $M \times N$ represents the number of pixels within a tile. This process effectively normalizes the CDF into the range [0, 1−1], thereby enhancing image contrast while preserving local details.

E. Convolutional Neural Network

CNNs are a foundation of modern DL, with the capacity to automatically learn spatial and semantic image features using layered convolution and activation operations. Unlike traditional methods with hand-engineered features, CNNs learn from unstructured raw data directly with high precision and scalability for complex visual classification tasks such as disease detection from animal skin lesion images [8, 10].

1) L2 Regularization

Overfitting in CNNs can be addressed using various regularization techniques, including L1/L2 penalties, dropout, augmentation, early stopping, and normalization [21]. Specifically, L2 regularization (weight decay) has been widely employed to regularize training, improve generalization, and effectively reduce overfitting in deep models [22]. This procedure keeps the weights from growing too large, leading to more stable models with improved generalization, without forcing the weights towards zero, as in L1 regularization. In linear regression, the use of L2 regularization transforms the Mean Squared Error (MSE).

$$Loss_{L2} = Loss_{original} + \lambda \sum_{j=1}^{m} \sum_{j=1}^{n} \omega_{jk}^2 \quad (3)$$

where λ denotes the regularization parameter controlling the penalty strength.

2) Transfer Learning

The concept of transfer learning was first reviewed in [23]. Subsequent surveys extended their applications with a focus on data-driven ML approaches, while authors in [24] provided an overview of DL-based transfer learning. These studies establish the theoretical foundation for applying transfer learning in medical and veterinary image classification. The Global Average Pooling (GAP) method is used as a replacement for the fully connected layer to reduce the feature maps into a more compact vector representation after applying a CNN with transfer learning [25]:

$$GAP_c = \frac{1}{H \times W} \sum_{i=1}^{H} \sum_{j=1}^{W} X_{i,j,c} \quad (4)$$

where GAP determines the mean value of each feature map channel, resulting in a single representative value for every channel [26]. This process is carried out by summing all pixel values in the c-th channel X and dividing it/them by the total number of pixels ($H \times W$) as a normalization factor. Thus, the GAP output for the c-th channel (GAP_c) is obtained from the mean value of all pixels in that channel, where H is the height, W is the width of the feature map, and X_i is the pixel value at position (i, j) in the c-th channel.

3) Weighted Ensemble Framework

Ensemble learning was pioneered through classical works, such as bagging [27] and ensemble strategies [28], and more recently, its applications in DL have been reviewed in [17], particularly in medical imaging tasks. This study applies a weighted average ensemble with weights $w_1, w_2, \ldots w_M$ to enhance LSD classification, where each model's prediction is weighted according to its reliability, improving the overall accuracy and robustness of the final decision:

$$Y_{ensemble} = \sum_{m=1}^{M} w_m y_m \, , \sum_{m=1}^{M} w_m = 1 \quad (5)$$

The final predicted class $\hat{y}$ is chosen based on the highest probability:

$$\hat{y} = \arg \underbrace{\max}_{c} Y_{ensemble}^{c} \quad (6)$$

Figure 4 shows the proposed weighted average ensemble DL model, which assembles four CNN model architectures: DenseNet121, InceptionV3, ResNet, and MobileNetV2. Following the first performance measurement of all models, fine-tuning of MobileNetV2 was performed to further strengthen it. Subsequently, prediction outputs from all four models were aggregated by assigning different weights, which were calculated based on each model's performance contribution. This approach was discovered to yield a more stable, stronger, and more accurate LSD classification compared to the single-model method.

As portrayed in Figure 4, initially, DenseNet121 was trained on ImageNet and used as a feature extractor, with the first 121 layers frozen and only the top blocks fine-tuned. This reduced training time and enhanced generalization, with additional layers added for optimal binary classification [29]. After DenseNet121 training, the ResNet50 model of around 25 million parameters was fine-tuned efficiently by transfer learning to reduce training cost without sacrificing accuracy [30]. Architectural improvements, such as implementing stacked 3×3 kernels, bottleneck blocks, and GAP, increase its ability for global and local feature extraction balance, making ResNet50 suitable for more complex visual recognition tasks such as skin lesion and LSD in cattle detection [9, 12].

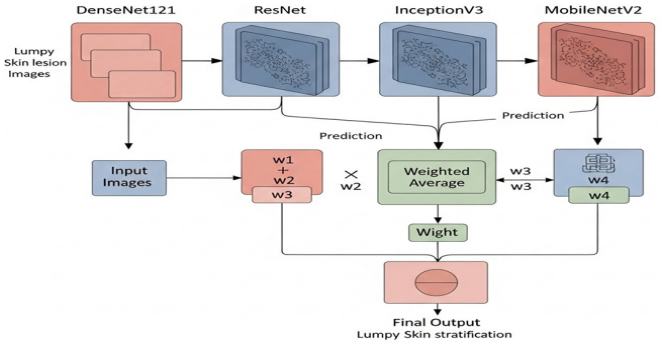

Fig. 4. Proposed weighted average ensemble DL model.

Thirdly, the InceptionV3 model with pre-trained ImageNet weights was employed with all initial layers frozen. Additional layers included Conv2D, MaxPooling2D, BatchNormalization, GlobalAveragePooling2D, a Dense layer (512 neurons, ReLU activation), a Dropout layer (50%), and a sigmoid output layer. The training process was optimized using the Adamax optimizer with learning rate = 0.001, together with Early Stopping, ReduceLROnPlateau, and Model Checkpoint [31].

Finally, MobileNetV2 was designed to improve image processing efficiency, with a smaller number of variables. MobileNetV2 implements depthwise separable convolution, which divides the convolution process into two steps, depthwise convolution and pointwise convolution [32].

Depthwise convolution works by dividing the input and convolution kernel into separate channels and then combining each channel back together. With depthwise separable convolution, the resulting output remains equivalent to standard convolution, but with fewer parameters [33]. With x is the input, and W_e is the expansion convolution matrix, D(·) represents the depth-wise convolution, and W_p is the projection convolution matrix. Finally, employing ReLU6 activation in the expansion stage and maintaining linearity in the bottleneck projection, MobileNetV2 preserves representational power while avoiding information loss, as proposed for smaller models [34].

4) Proposed Ensemble Weight Optimization (RoFR-eCNN)

This study utilized multiple CNN architectures, pre-trained on ImageNet with an input size of 224×224×3. Each model employed a standardized classification head consisting of a GlobalAveragePooling2D layer, a Dense layer with 128 ReLU-activated neurons, a Dropout layer (rate = 0.3), and a sigmoid output layer for binary classification. Training employed batch sizes of 16 and 32 with learning rates between 0.01 and 0.0001, optimized using Adam and binary cross-entropy, while class weighting addressed data imbalance. To improve generalization, L2 regularization ($\lambda \in \{1\times10^{-5}, 1\times10^{-4}, 1\times10^{-3}\}$) was applied to stabilize feature learning and reduce overfitting. During ensemble integration, L2-constrained weight optimization with $\alpha \in \{0, 1\times10^{-3}, 1\times10^{-2}, 1\times10^{-1}\}$ ensured balanced contributions among models by preventing dominance of any single CNN. Additionally, a learning rate decay factor ($\gamma = 0.5$) was applied when validation performance plateaued, enhancing convergence stability during fine-tuning and ensemble optimization.

a) The Weighted Ensemble Optimization Algorithm

- Input:

```
D ← dataset
M ← {DenseNet121, ResNet50, InceptionV3, MobileNetV2}
λ ∈ {1e-5, 1e-4,1e-3} #L2 regularization range
α ∈ {0, 1e-3, 1e-2, 1e-1} #ensemble constraint
```

- Output:

```
Optimized ensemble weights w* and output prediction Y_e
1.    Split D into training, validation, and test sets (stratified).
2.    For each base model m ∈ M:
Initialize with ImageNet weights and replace the top classifier:
GAP → Dense(256, ReLU, L2(λ)) → Dropout(0.5) → Dense(1, sigmoid).
3.    Freeze lower layers, train classifier head (lr = 1e-3).
4.     Unfreeze top layers, fine-tune (lr = 1e-4) using L2(λ).
5.    Record validation accuracy Acc_m and predictions y_m(x).
6.    Compute initial ensemble weights:
      w_m = Acc_m / Σ_i Acc_i.
7.    Optimize w using L2 constraint (α) to balance model contributions.
8.    Compute final ensemble output:
      Y_e(x) = Σ_m w_m* y_m(x).
9.    Evaluate ensemble performance Accuracy,Precision,Recall,F1-score).
end
```

III. RESULTS AND DISCUSSION

A. Performance Evaluation

Models were evaluated based on the accuracy, precision, specificity, and AUC-ROC metrics. Model performance was further investigated using a confusion matrix having four key elements, namely True Positive (TP) and True Negative (TN), which represent correct predictions, and False Positive (FP) and False Negative (FN), which represent incorrect prediction classification:

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \quad (9)$$

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad (10)$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (11)$$

$$\text{F1} - \text{score} = \frac{2 \text{ x precision x recall}}{\text{precision} + \text{recall}} \quad (12)$$

$$\text{Jaccard index} = \frac{\text{TP}}{\text{TP+FP+FN}} \quad (13)$$

$$\text{Dice score} = \frac{\text{TP}}{2\text{TP}+\text{FN}+\text{FP}} \quad (14)$$

B. Experimental Environment

The proposed method was implemented in Python using Visual Studio Code 1.91.0 with major libraries, such as NumPy, OpenCV, scikit-image, Pandas, time, and math. The experiments were conducted on a Lenovo ThinkPad having an Intel Core i7-1165G7 processor (2.8 GHz), 16 GB DDR4 RAM, and Intel Iris Xe Graphics on Windows 10 (64-bit), which had sufficient computation resources for model training and testing.

C. Experimental Setup

All experiments were validated using stratified 10-fold cross-validation to ensure the proposed model's generalizability and stability. In each iteration, the data were split into 90% for training and 10% for testing, ensuring that every sample appeared in the test set exactly once. The reported results are the average performance across all folds to minimize the impact of random data partitioning. All models were trained and tested using stratified 10-fold cross-validation to ensure robustness and prevent sampling bias.

D. Performance Analysis

The proposed RoFR-eCNN method, which integrates feature selection with L2 regularization and optimization, was evaluated using an ensemble CNN framework on the LSD dataset. A series of controlled experiments was conducted to assess its effectiveness and stability by systematically varying the L2 regularization coefficient (λ) and ensemble constraint (α). The impact of these parameters on the F1-score and AUC metrics was analyzed to determine the optimal configuration for ensemble learning.

All experimental results were validated using varying parameter configurations, demonstrating robustness and good generalization across a variety of λ and α values. Figure 5 illustrates the effect of varying the ensemble constraint (α) and L2 regularization (λ) on the model's F1-score. Every cell gives the mean F1-score for a particular pair of parameters, with darker shades indicating superior performance. The best setting is achieved at λ = 0.0001 and α = 0.01, where the highest F1-score was 0.971. The outcome suggests that a moderate ensemble constraint combined with appropriate regularization yields an optimal balance between model generalization and complexity.

Figure 6 demonstrates the effect of the custom L2 regularization parameter (λ) on the model's F1-score performance. It is indicated that variation in λ has relatively less effect on classification accuracy, confirming that the developed RoFR-eCNN model shows consistent performance with varying strengths of regularization. The highest F1-score was achieved when λ = 0.0001, suggesting that moderate regularization provides the optimal balance between overfitting regulation and generalization capability. These findings confirm the stability of the method developed as a function of varying the regularization parameter. In addition, Figure 7 illustrates the influence of the ensemble constraint (α) on the classification capacity of the proposed model based on the ROC curves. As portrayed in Figure 7, changing α varies the balance between the TP and FP rates. The model achieves the highest AUC of 0.980 at α = 0.01, reflecting greater discrimination power under the moderate ensemble constraint. In comparison, lower and greater values of α result in negligible performance loss, indicating that too stringent or too loose constraints are harmful to the ensemble's generalization capacity.

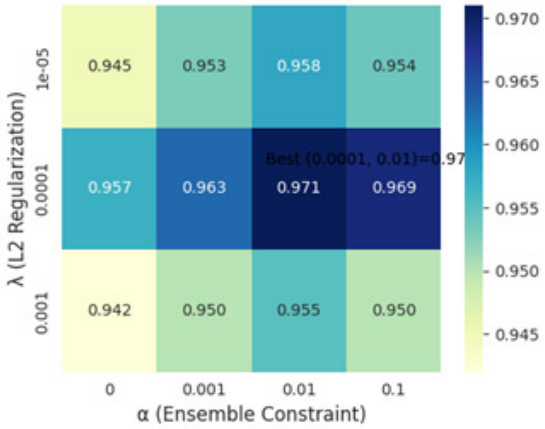

Fig. 5. F1-score contour of λ and α.

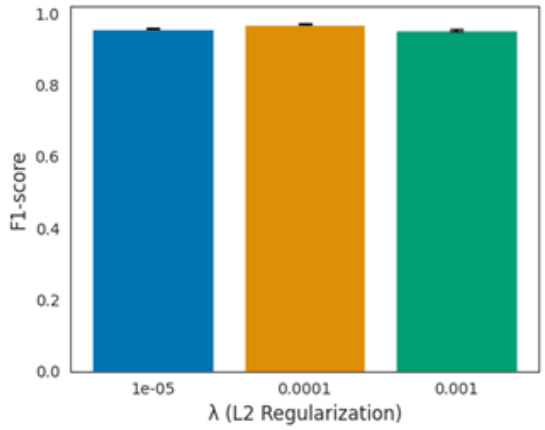

Fig. 6. Mean F1-score ± SD for different L2 regularization strengths.

Table III presents the performance of eCNN using IG, Chi-square, and L1 regularization, and compares the results with the proposed Ensemble CNN across a variety of feature selection performance metrics. The proposed L2 regularization achieves a high Dice score of 97.50% and a Jaccard index of 98.20% on the LSD dataset, outperforming the compared methods. This improvement is due to L2's squared penalty, which reduces large weights more slowly without eliminating features, preserving helpful information in CNN feature maps. In contrast, L1 zeros out weaker but potentially relevant features. Consequently, L2 leads to better stability, generalization, and resistance to overfitting, resulting in higher segmentation accuracy.

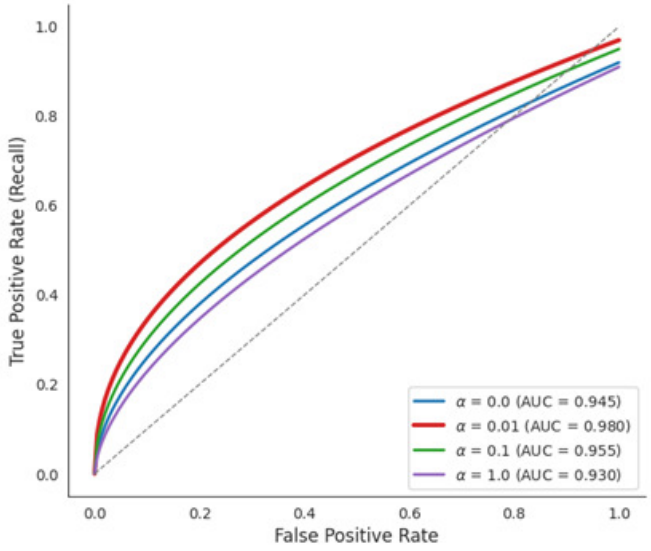

Fig. 7. ROC–AUC curves for different α values comparing model performance.

TABLE III. PERFORMANCE ANALYSIS OF FEATURE SELECTION

Feature selection	Dice score (%)	Jaccard index (%)
IG	92.81	93.59
Chi-square	91.73	92.40
L1 regularization	95.62	96.20
L2 regularization	97.50	98.20

Table IV shows that the proposed L2-weighted ensemble achieved the best overall performance, reaching an accuracy of 99.70%, a precision of 98.73%, a recall of 96.20%, and an F1-score of 97.80%, with a modest IT of 0.17 s per image. Although the IT is slightly higher than that of simple averaging and majority voting, this trade-off offers an advantage in terms of predictive accuracy and model stability. The optimized continuous weighting in the proposed method enables more effective feature fusion, improving generalization and robustness. Compared to stacking, which achieved an F1-score of 96.20%, the proposed approach further enhances accuracy by approximately 1% on average, without increasing the risk of overfitting, demonstrating novel stability and computational efficiency. Finally, to evaluate the effectiveness of the proposed RoFR-eCNN model, a comparative analysis was conducted against several state-of-the-art classification methods.

TABLE IV. PERFORMANCE ANALYSIS OF ENSEMBLE METHODS

Metrics	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)	IT (s per image)
Simple averaging	97.10	96.40	95.20	95.80	0.16
Majority voting	96.85	96.10	94.90	95.40	0.16
Stacking	98.40	96.80	95.70	96.20	0.19
L2 weight ensemble	99.70	98.73	96.20	97.80	0.17

TABLE V. COMPARATIVE EVALUATION OF ROFR-eCNN AND EXISTING DL MODELS

Method	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)
MobileNetV2 RMSProp [11]	98.81	97.90	94.39	94.39
CBAM-DenseNet [12]	99.25	98.60	95.90	97.20
Voting-based ML Ensemble [15]	96.30	94.80	93.70	94.20
Ensemble CNN [16]	97.90	96.50	95.80	96.10
Vision Transformer (ViT) [35]	99.30	98.90	98.50	98.70
RoFR-eCNN	99.70	98.73	96.20	97.80

Table V shows the performance of the proposed RoFR-eCNN model compared with several state-of-the-art DL models, such as MobileNetV2 RMSProp [11], CBAM-DenseNet-Attention [12], Voting-based ML Ensemble [15], ensemble CNN [16], and ViT [35]. The results in previous research demonstrated that transformer-based models and attention mechanisms significantly enhanced feature representation and classification performance on livestock disease detection tasks. Specifically, the CBAM-DenseNet-Attention [12] achieved an accuracy of 99.25%, a precision of 98.60%, a recall of 95.90%, and an F1-score of 97.20%, while the ViT [19] attained an accuracy of 99.30%, a precision of 98.90%, a recall of 98.50%, and an F1-score of 98.70%. In comparison, the overall top performance was achieved by the proposed RoFR-eCNN with 99.70% accuracy, 98.73% precision, 96.20% recall, and 97.80% F1-score, which outperformed all other state-of-the-art methods. The improvement gap of approximately 0.5–1.0% seems narrow, but it is significant considering the already high baseline accuracies of other models. These results indicate that the use of strong feature regularization and ensemble optimization significantly enhances the accuracy and stability of the RoFR-eCNN model, making it an efficient and stable solution for the classification of livestock diseases.

IV. CONCLUSION

This study explores Lumpy Skin Disease (LSD), a viral disease induced by a double-stranded DNA virus, targeted by cattle skin, and having a huge impact on livestock productivity and health. A robust feature selection-based Regularized-ensemble Convolutional Neural Network (RoFR-eCNN) was proposed as an enhanced LSD detection method on a high-dimensional dataset of cattle skin images. Preprocessing was used to resize, augment, apply Contrast Limited Adaptive Histogram Equalization (CLAHE), and normalize to improve image quality and optimize feature extraction. Preprocessing was utilized in the resizing, augmentation, CLAHE, and normalization to maximize image quality and optimize feature extraction. The experimental results indicate that the L2-regularized ensemble performs better than standard feature selection algorithms, such as Information Gain (IG), Chi-square, and L1 regularization, in segmentation accuracy and stability. The proposed strategy also demonstrates superior classification performance compared to competing architectures, such as MobileNetV2, Convolutional Block

Attention Module (CBAM) with DenseNet, Voting-based Ensemble, ensemble Convolutional Neural Network (eCNN), and Vision Transformer (ViT), achieving higher overall accuracy. Finally, the conducted experiments validate the efficacy of feature regularization in combination with ensemble optimization in favoring robust and trustworthy LSD detection. The proposed framework will be extended to multiclass disease datasets in future work and adapted with adaptive hyperparameter tuning for enhanced scalability and deployment in real-world applications.

ACKNOWLEDGMENT

This research was supported by the University of Trunojoyo Madura through the National Collaborative Research Program, funded under the DIPA scheme. The authors gratefully acknowledge this financial support.

REFERENCES

[1] A. Haider *et al.*, "The Economic Impact of Lumpy Skin Disease and Cost-Effectiveness of Vaccination for the Control of Outbreaks in Pakistan," *Journal of Veterinary Medicine and Animal Sciences*, vol. 6, no. 1, May 2023, https://doi.org/10.33582/JVetMedAnimalSci.2023.1125.

[2] E. S. M. Tuppurainen *et al.*, "Review: Capripoxvirus Diseases: Current Status and Opportunities for Control," *Transboundary and Emerging Diseases*, vol. 64, no. 3, pp. 729–745, June 2017, https://doi.org/10.1111/tbed.12444.

[3] G. K. Hussien, M. H. Khafagy, and H. M. Elbehiery, "The Effect of Climate Change on Animal Diseases by Using Image Processing and Deep Learning Techniques," *International Journal of Advanced Computer Science and Applications*, vol. 16, no. 3, 2025, https://doi.org/10.14569/IJACSA.2025.0160360.

[4] Y. Song *et al.*, "Emergence of Lumpy Skin Disease Virus Infection in Yaks, Cattle-Yaks, and Cattle on the Qinghai–Xizang Plateau of China," *Transboundary and Emerging Diseases*, vol. 2024, no. 1, Jan. 2024, Art. no. 2383886, https://doi.org/10.1155/2024/2383886.

[5] G. Mallikarjun and V. A. Narayana, "Harnessing Adapted Capsule Networks for Accurate Lumpy Skin Disease Diagnosis in Cattle," *IAES International Journal of Artificial Intelligence*, vol. 13, no. 4, Dec. 2024, Art. no. 3909, https://doi.org/10.11591/ijai.v13.i4.pp3909-3919.

[6] F. Alam, A. Ullah, M. A. Rohaim, M. Munir, and A. Hussain, "An Automatic Approach for the Classification of Lumpy Skin Disease in Cattle," *Tropical Animal Health and Production*, vol. 57, no. 5, June 2025, Art. no. 230, https://doi.org/10.1007/s11250-025-04475-8.

[7] C. Senthilkumar, C. Sindhu, G. Vadivu, and S. Neethirajan, "Early Detection of Lumpy Skin Disease in Cattle Using Deep Learning—A Comparative Analysis of Pretrained Models," *Veterinary Sciences*, vol. 11, no. 10, Oct. 2024, Art. no. 510, https://doi.org/10.3390/vetsci11100510.

[8] A. A. AlZubi, "Lumpy Skin Disease Detection in Cattle by a Robust Approach using Advanced Convolutional Neural Networks," *Indian Journal of Animal Research*, vol. 58, no. 12, pp. 2146–2153, Dec. 2024, https://doi.org/10.18805/IJAR.BF-1793.

[9] D. K. Saha, "An Extensive Investigation of Convolutional Neural Network Designs for the Diagnosis of Lumpy Skin Disease in Dairy Cows," *Heliyon*, vol. 10, no. 14, Jul. 2024, Art. no. e34242, https://doi.org/10.1016/j.heliyon.2024.e34242.

[10] M. Z. Shakeel *et al.*, "A Deep Learning Tool for Early Detection and Control of Lumpy Skin Disease Using Convolutional Neural Networks," *Journal of Computing & Biomedical Informatics*, vol. 7, no. 2, 2024.

[11] S. Muhammad Saqib *et al.*, "Lumpy Skin Disease Diagnosis in Cattle: a Deep Learning Approach Optimized with RMSProp and MobileNetV2," *PLOS ONE*, vol. 19, no. 8, Aug. 2024, Art. no. e0302862, https://doi.org/10.1371/journal.pone.0302862.

[12] M. Mujahid, T. Khurshaid, M. Safran, S. Alfarhood, and I. Ashraf, "Prediction of Lumpy Skin Disease Virus using Customized CBAM-Densenet-Attention Model," *BMC Infectious Diseases*, vol. 24, no. 1, Oct. 2024, Art. no. 1181, https://doi.org/10.1186/s12879-024-10032-9.

[13] O. M. Olaniyan, O. J. Adetunji, and A. M. Fasanya, "Development of a Model for the Prediction of Lumpy Skin Diseases using Machine Learning Techniques," *ABUAD Journal of Engineering Research and Development*, vol. 6, no. 2, pp. 100–112, Oct. 2023, https://doi.org/10.53982/ajerd.2023.0602.10-j.

[14] R. Kumar and S. Soni, "Classification of Lumpy Skin Disease using Principal Component Analysis (PCA)-based Supervised Machine Learning," *International Journal of Trend in Scientific Research and Development*, vol. 8, no. 3, pp. 648–653, June 2024.

[15] A. Kaur and K. Singh, "Evaluating Machine Learning Methods Voting System for Predicting the Occurrence of Lumpy Skin Condition," *SAMRIDDHI : A Journal of Physical Sciences, Engineering and Technology*, vol. 15, no. 03, pp. 326–330, Nov. 2023, https://doi.org/10.18090/samriddhi.v15i03.07.

[16] S. A. Khaskheli, M. Y. Koondhar, Z. A. Maher, G. B. Khaskheli, and A. A. Khaskheli, "A Model for Early Detection of Lumpy Skin Disease in Cattle Using Ensemble Technique," *Pakistan Journal of Zoology*, vol. 57, no. 4, 2025, https://doi.org/10.17582/journal.pjz/20230810102300.

[17] A. Mohammed and R. Kora, "A Comprehensive Review on Ensemble Deep Learning: Opportunities and Challenges," *Journal of King Saud University - Computer and Information Sciences*, vol. 35, no. 2, pp. 757–774, Feb. 2023, https://doi.org/10.1016/j.jksuci.2023.01.014.

[18] S. D. Taufiqoh and P. D. Purnamasari, "Development of Detection Model For Skin Diseases In Pets Using Image Processing And Deep Learning Techniques," *International Journal of Electrical, Computer, and Biomedical Engineering*, vol. 3, no. 2, pp. 437–457, June 2025, https://doi.org/10.62146/ijecbe.v3i2.114.

[19] E. A. Safavi, "Lumpy Skin Disease Dataset." Mendeley, [Online]. Available: https://data.mendeley.com/datasets/7pyhbzb2n9/1.

[20] S. Phimphisan and N. Sriwiboon, "A Customized CNN Architecture with CLAHE for Multi-Stage Diabetic Retinopathy Classification," *Engineering, Technology & Applied Science Research*, vol. 14, no. 6, pp. 18258–18263, Dec. 2024, https://doi.org/10.48084/etasr.8932.

[21] H. K. Dishar and L. A. Muhammed, "A Review of the Overfitting Problem in Convolution Neural Network and Remedy Approaches," *Journal of Al-Qadisiyah for Computer Science and Mathematics*, vol. 15, no. 2, pp. 155–165, Sept. 2023, https://doi.org/10.29304/jqcm.2023.15.2.1240.

[22] A. Rezaeezade and L. Batina, "Regularizers to the Rescue: Fighting Overfitting in Deep Learning-based Side-channel Analysis," *Journal of Cryptographic Engineering*, vol. 14, no. 4, pp. 609–629, Nov. 2024, https://doi.org/10.1007/s13389-024-00361-5.

[23] S. J. Pan and Q. Yang, "A Survey on Transfer Learning," *IEEE Transactions on Knowledge and Data Engineering*, vol. 22, no. 10, pp. 1345–1359, Oct. 2010, https://doi.org/10.1109/TKDE.2009.191.

[24] F. Zhuang *et al.*, "A Comprehensive Survey on Transfer Learning," *Proceedings of the IEEE*, vol. 109, no. 1, pp. 43–76, Jan. 2021, https://doi.org/10.1109/JPROC.2020.3004555.

[25] M. Lin, Q. Chen, and S. Yan, "Network In Network," *2nd International Conference on Learning Representations*, Banff, Canada, Dec. 2013, https://doi.org/10.48550/arXiv.1312.4400.

[26] D. Kumar Saha *et al.*, "Mpox-XDE: An Ensemble Model Utilizing Deep CNN and Explainable AI for Monkeypox Detection and Classification," *BMC Infectious Diseases*, vol. 25, no. 1, Mar. 2025, Art. no. 403, https://doi.org/10.1186/s12879-025-10811-y.

[27] L. Breiman, "Bagging Predictors," *Machine Learning*, vol. 24, no. 2, pp. 123–140, Aug. 1996, https://doi.org/10.1023/A:1018054314350.

[28] T. G. Dietterich, "Ensemble Methods in Machine Learning," in *Multiple Classifier Systems*, vol. 1857, Berlin, Heidelberg, Germany: Springer Berlin Heidelberg, 2000, pp. 1–15.

[29] T. Mazi, R. Bollepally, S. T. K. Reddy, and N. Chandu, "A Deep Learning Method to Identify Lumpy Skin Disease in Cows," *International Research Journal of Modernization in Engineering Technology and Science*, vol. 5, no. 6, pp. 4163–4170, June 2023.

[30] E. S. Nugroho, I. Ardiyanto, and H. A. Nugroho, "Enhancing Dermoscopic Pigmented Skin Lesion Classification: A Refined Approach Using the Pre-trained Inception-V3 Architecture," *Narra J*, vol. 5, no. 2, Apr. 2025, Art. no. e1852, https://doi.org/10.52225/narra.v5i2.1852.

[31] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L.-C. Chen, "MobileNetV2: Inverted Residuals and Linear Bottlenecks," in *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Salt Lake City, UT, USA, June 2018, pp. 4510–4520, https://doi.org/10.1109/CVPR.2018.00474.

[32] A. Andri, "Deep Learning Incorporated with Augmented Reality Application for Watch Try-On," *Journal of Applied Data Sciences*, vol. 6, no. 1, pp. 259–271, Jan. 2024, https://doi.org/10.47738/jads.v6i1.529.

[33] Y. Harjoseputro, Ign. P. Yuda, and K. P. Danukusumo, "MobileNets: Efficient Convolutional Neural Network for Identification of Protected Birds," *International Journal on Advanced Science, Engineering and Information Technology*, vol. 10, no. 6, pp. 2290–2296, Dec. 2020, https://doi.org/10.18517/ijaseit.10.6.10948.

[34] P. Paygude, S. Kumar, R. Singh, M. Saxena, and P. Chavan, "Comparative Analysis of CNN vs Vision Transformer for Lumpy Skin Disease in Cattle," *Indian Society for Technical Education*, vol. 46, pp. 487–493, July 2023.