



A Deep Learning-Based Integrated Smart Dairy Farming System for Cattle Health Monitoring and Milk Quality Analysis

Pratik Shingade¹, Adhik Sarak²,
Abhishek Walkoli³, Lucky Ragde⁴,
Roshani Raut⁵, Radha Deoghare⁶

^{1,2,3,4,5,6} Pimpri Chinchwad College of Engineering Pune, India

pratik.shingade23@pccoepune.org, adhik.sarak23@pccoepune.org*

, abhishek.walkoli23@pccoepune.org*,

* lucky.ragde23@pccoepune.org*, roshani.raut@pccoepune.org*,

radha.deoghare@pccoepune.org*

Abstract. The modern dairy farming is shifting more towards intelligent and data-driven systems to improve productivity, animal welfare, and milk quality measurement. Nevertheless, conventional methods of monitoring are also based on manual observation leading to the late diagnosis of the disease and the lack of a consistent quality assessment. This research paper presents a proposed smart dairy farming system that integrates the deep learning and machine learning algorithms to create a real-time cattle health monitoring and milk quality predictive tool. Evaluation of various supervised learning models was done using a structured dataset comprising of physicochemical and sensory attributes. The highest precision, recall and F1-score was obtained with the Random Forest classifier with the highest accuracy of 95, compared to the K-Nearest Neighbors, Naive Bayes and Support Vector Machine classifiers. To detect cattle diseases, the EfficientNetV2 deep learning architecture was used in order to classify diseases like Lumpy Skin Disease, Bovine infections and Foot-and-Mouth Disease based on images. The model performed better as it was able to extract features efficiently and generalize features effectively. The proposed system combines the multi-domain data and offers a scalable, automated and reliable solution to the modern day management and decision support of a dairy farm.

Keywords: *Smart dairy farming, cattle health monitoring, Deep learning, EfficientNetV2, Random Forest classifier, precision livestock farming.*

1 Introduction

Industries within the global agricultural sector has been the dairy industry, which has been playing a significant role in food security, rural economy and nutritional supply chains. As the demand of dairy products has been increased, there is a growing need of smart and automated dairy farm management systems. The contemporary dairy industry encounters a number of challenges such as cattle health management, cattle feeding management, cattle productivity management and milk quality management which in turn have a direct impact on financial sustainability.

Conventional agricultural methods are based on human observation, veterinary skills and laboratory analysis. These are labor intensive, time consuming methods which are mostly subjective. As Krishna et al. [1] emphasized, such methods result in the late detection of diseases and the high operational costs. Karthikeyan et al. [2] also highlighted that large-scale dairy farms are hugely affected because of the late detection of abnormalities which lead to low milk production and economic losses.

The importance of cattle health monitoring is explained by the fact that it is directly

related to milk production and the profitability of farms. Mastitis, Lumpy Skin Disease (LSD), Bovine infections, and Foot-and-Mouth Disease (FMD) have a negative impact on animal welfare and milk quality. According to Bhardwaj et al. [3], early symptoms of these diseases are often subtle and not easily detected manually, thus requiring automated diagnostic solutions.

Health monitoring systems based on machine learning have proven to be highly promising when it comes to early detection of disease with the help of the physiological and behavioral information. Jia et al. [6] demonstrated that feeding, standing, and resting behaviors of cattle could be effectively categorized to predict the health conditions with high accuracy. These results suggest that automated systems have the potential to dramatically improve disease detection.

Moreover, feeding behavior is an important factor in the livestock production and welfare. Early signs of stress, metabolic disorders, and disease conditions are considered to be changes in the feeding frequency and duration, which is discussed by Suparyati et al. [9]. The Existing research having following limitations:

1. Lack of integrated systems combining cattle disease detection and milk quality prediction
 2. Limited use of deep learning for real-time disease diagnosis
 3. Absence of cross-domain data utilization (health ↔ milk quality)
 4. Minimal deployment-ready solutions for farmers
- The contributions of this paper can be listed as follows:

1. Development of an integrated smart dairy system combining disease detection and milk quality prediction
2. Implementation of EfficientNetV2 for image-based cattle disease classification
3. Comparative analysis of multiple ML models for milk quality prediction
4. Identification of Random Forest as the optimal classifier (95% accuracy)
5. Design of a scalable architecture suitable for real-world deployment

2 Literature Review

The development of deep learning technologies has made it possible to dramatically advance the development of artificial intelligence and machine learning, and their application in dairy farming. Researchers are enjoying attention on automated systems to monitor cattle health, feeding behaviors, and milk quality as part of precision livestock farming systems. Initial works in the field were mostly based on physiological data and environmental conditions along with machine learning methods to detect early diseases. Gokul Krishna et al. [1] suggested a machine learning-based diagnostic system, which uses sensor-derived parameters to detect diseases in cows at their early stages. Their work shows that data-driven approaches can be used to enhance prediction accuracy and minimize diagnostic latency. The study, furthermore points out the shortcomings of traditional manual monitoring systems and the necessity of automated decision support systems in the contemporary farm management. Subsequent studies have resulted in the design of a few machine learning-based livestock monitoring systems, which can process clinical data, physiological data, and behavioral data to detect and prevent diseases. Karthikeyan et al. [2] made a contribution to this field, as they developed a machine learning framework to predict diseases, which supports the significance of smart monitoring systems in livestock management. Health monitoring of behavior is an important line of research that has become a critical direction. In a different study, Jia et al. [6] designed a hybrid solution based on inertial sensor data and Random Forest algorithms and correlation analysis. In their study, they found strong correlations between behavioral patterns of a specific kind and the health conditions of the cattle. The model was found to have high classification performance in determining

activities like feeding, standing, resting, and ruminating, which are considered important indicators of health in cattle. The research also showed that such combined states of behavior as feeding-standing and resting-lying may be reliable signs of animal health. It also demonstrated that the methods of feature selection used in correlation-based methods may lead to improved classification performance without raising the complexity of the model. These results show the significance of behavioral analytics as the practical tool to detect diabetes in the early stages and to monitor the welfare of livestock. Simultaneously, computer vision and deep learning technologies have revolutionized cattle monitoring systems. Architectures that are built on EfficientNetV2 have demonstrated high performance in the detection of visual symptoms of diseases. Mhapsekar et al. [4] showed that deep learning models are more effective than the traditional machine learning methods in identifying meaningful features of the complex image data. In contrast to traditional systems that rely on manual feature engineering, deep learning models do not need any form of manual feature engineering and can automatically detect the symptoms, e.g. skin lesions, nodules and abnormalities. The comparative analysis of lightweight and optimized deep neural network architectures shows that they are more likely to succeed in real-life farming settings because of the enhanced generalization abilities. These results justify using deep learning methods in livestock health monitoring systems. Another significant issue of dairy farming is the feeding behavior monitoring as it directly affects productivity and metabolic health. Li et al. [5] introduced a computer vision system based on deep learning to analyze behavior and demonstrated that monitoring based on cameras could be able to identify feeding patterns and recognize anomalies. Nevertheless, their method also emphasizes the drawbacks of the traditional sensor-based systems like RFID that are associated with complex infrastructure and constant maintenance. The findings of these studies suggest that the visual monitoring systems based on AI offer scalable and efficient alternatives to real-time livestock monitoring. These systems can be used to help farmers optimize feeding schedules and identify early signs of stress or disease. Further research also highlights the significance of automated analytics in enhancing the performance of dairy farms. Suparyati et al. [9] wrote about the machine learning techniques of predicting livestock diseases, stating that there are multiple advantages of using data-driven decision-making in managing farms. The other urgent research field is the evaluation of milk quality with the help of artificial intelligence. It is common that traditional machine learning methods cannot efficiently handle the high-dimensional and nonlinear milk data and thus requires a lot of preprocessing and feature engineering. Kamilaris and Prenafeta-Boldu [7] pointed to the increased significance of deep learning when processing highly complex agricultural data. Deep learning models have been shown to have higher ability to capture intricate relationships in milk quality data. Mhapsekar et al. [4] proposed a Hybrid Blended Deep Learning (HyBDL) that combines EfficientNetV2, CNN-LSTM, and CNN-GRU architecture. Their methodology has attained higher classification quality and strength in detecting milk adulteration. The paper also emphasized that hybrid deep learning models can be very useful in learning hierarchical and time-dependent feature representations, eliminating the reliance on domain-specific feature engineering. Moreover, the ensemble deep learning strategies have been reported to enhance the stability of prediction, reproducibility and hyper-generalization performance as has been supported by Afshari-Safavi et al. [8]. Nevertheless, the majority of current studies are dedicated to single areas like cattle health monitoring, feeding behavior analysis or milk quality prediction. What is lacking is the integration of such systems that will bring together such aspects into a single system. In current studies, these components are often viewed as autonomous issues instead of processes that are interconnected. Feeding behavior, physiological health, and milk quality are interdependent variables that in combination are known to affect overall productivity of the farm. Thus, the important research requirement is the development of an integrated analytical model, which will capture these relationships. Driven by these gaps, the current study suggests an integrated smart dairy farming solution, which consists of a combination of deep learning-based cattle disease detection and machine learning-based milk quality prediction. The proposed framework will enhance the accuracy, scalability, and automation of the prediction and minimize the complexity of the infrastructure.

3 Methodology and Dataset

3.1 Dataset Information

3.1.1 Milk Quality Dataset

The milk quality data comprise 10,001 samples, each of which depicts the physicochemical and sensory characteristics of milk. The important attributes in the dataset are:

- pH
- Temperature
- Taste
- Odor
- Fat
- Turbidity
- Color

Milk quality grade is the target variable in this data set and it is divided into Low, Medium, and High categories. They are very common parameters used to determine the freshness and quality of milk in general.

The data gives an organized form that can be used by the supervised machine learning models and offers the ability to accurately classify the milk quality based on different conditions.

3.1.2.Cattle Disease Image Dataset.

In cattle health monitoring, an image based dataset was employed, and this consisted of labeled images depicting the various health conditions of cattle. In the dataset, there are the following classes:

- a. Bovine Disease
- b. Foot and Mouth Disease (FMD)
- c. Lumpy Skin Disease (LSD)
- d. Healthy

The milk quality data were collected in the form of publicly available machine learning repositories, and the cattle disease data were gathered in the form of open-source agricultural datasets and annotated livestock disease image collections.

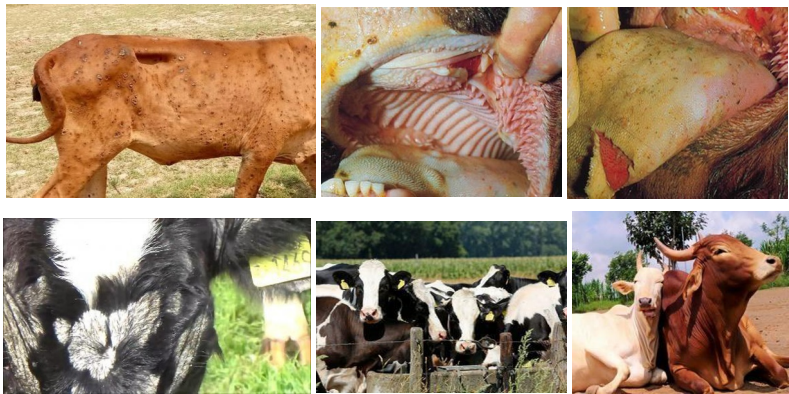


Fig.1. Images from the Dataset

3.2 Data Preprocessing

The first and the most essential step to be taken in order to guarantee better stability and accuracy of the cattle disease detection model is image preprocessing.

The preprocessing pipeline entails:

- Image normalization to standard pixel intensity distribution.
- Downsizing to achieve standard input size.
- The methods of data augmentation, including:
 - Rotation
 - Reflection (flipping)
 - Zooming
 - Variation in brightness and intensity.

These techniques help:

- Increase dataset diversity
- Prevent overfitting
- Improve model generalization

Consequently, EfficientNetV2 is efficient in picking discriminative features, and the model can distinguish very fast between healthy and diseased cattle.

3.3 Raw and Engineered Features

Table 1. Raw and the Engineered Features

Module	Raw Features	Engineered Features	Description
Milk Quality Prediction	pH, Temperature, Taste, Odor, Fat, Turbidity, Colour	- Normalized numerical characteristics. - Encoded categorical variables - Interaction of features (e.g., pH-Temperature, Fat-	Raw physicochemical and sensory features are scaled and converted into numerical values to enhance learning of the model.

		Turbidity)	
Cattle Disease Detection	Raw cattle images (RGB)	- Normalized pixel values - Image enhancement (rotation, inverting, zooming, brightness)	- Deep features with EfficientNetV2 Image data is processed and refined to enhance features extraction and classification accuracy.

3.4 Model Architecture

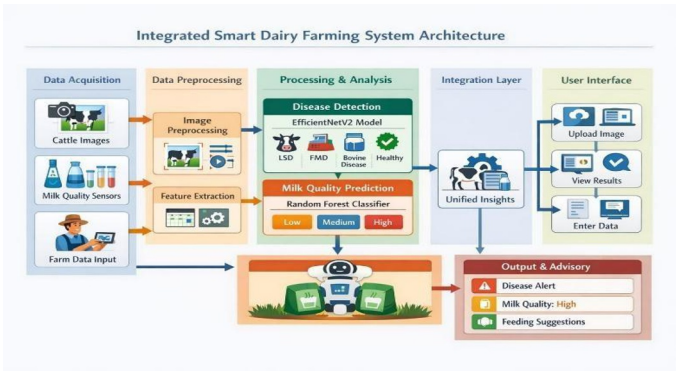


Fig.2. Architecture of System

The given system is offered as one comprehensive smart dairy farming system which will integrate both cattle disease detection and milk quality prediction. The first stage involves collecting data on various sources, such as milk images of cattle and well-organized parameters of milk quality: pH, temperature, odor, fat, and turbidity. The data is preprocessed in the form of image data normalization and augmentation, and structured data optimization in terms of feature extraction methods.

To detect cattle diseases, a deep EfficientNetV2 is used to do so. learning architecture is taken. The suggested architecture is illustrated in Figure 2. EfficientNetV2 makes use of an optimized. convolutional blocks made to be efficient in feature extraction and enhanced performance of training. Hierarchical is captured in the network. patterns of cattle images, which allow identifying them successfully. of characters of disease. The architecture incorporates. non-linear activation functions to convolutional operations. acquire intricate representations relating to various cattle. health conditions. A fully employed classification stage is used in the last classification stage. followed by a multi-class softmax layer. disease prediction. To analyze milk quality, it is possible to use several machine learning classifiers. evaluated. Random According to comparative performance. The reason to use Forest (RF) classifier is because it has a higher accuracy. robustness. Random Forest uses ensemble approach to learning. whereby there are several decision trees that make a contribution to the. classification result, the improved generalization ability and minimizing overfitting. Besides that, the system has a chatbot-based feeding. help module that is created to offer interactive instructions. suggestions regarding the feeding and nutrition of cattle. Management

3.5 Modelling

Step 1: The use of EfficientNetV2 model used in. As an input to the system, photos of cattle are offered. weights to empower effective feature learning.

Step 2: The information goes through picture preprocessing methods like as image augmentation, image normalization and image scaling in. an effort to find out the training needs. The techniques enhance the extrapolation of the model and the uniformity. of the data.

Step 3: to identify patterns of discriminations that are associated with different disease. the visual characteristics of the states, which is done via the EfficientNetV2 classifier. cattle.

Step 4: Techniques of regularization are used to deter. overfitting and enhance capacity of the models to generalize. The ensemble method of the random forest makes sure that the approach is durable. the models are in the analysis of milk quality, and the Dropout method is efficacynetv2.0 used in the EfficientNetV2 model in disease classification.

Step 5: Data is divided into subsets to be used as both testing and training. The results of the tests are put aside to be evaluated in terms of performance. Model learning is done using training data.

Step 6: The performance of the EfficientNetV2 model is tested using cattle images, Random Forest classifier is tested by using parameters of milk quality.

3.6 EfficientNetV2 Classifier:

The model is an expansion on convolutional operations that can automatically learn hierarchical representations of features on image data. It uses optimized convolutional blocks which make it extract visual patterns related to cattle diseases, including skin abnormalities, lesions and inflammation.

Mathematically, the convolutional action within the network can be denoted by as:

$$Y_{i,j,k} = \sum_{m,n,c} X_{i+m,j+n,c} \cdot W_{m,n,c,k}$$

where X represents the input image, W denotes the learned kernel weights, and Y is the resulting feature map. Non-linear activation is applied using the Swish function:

$$\text{Swish}(x) = x \cdot \text{sigmoid}(x)$$

EfficientNetV2 also benefits from compound scaling, which uniformly balances network depth, width, and resolution:

$$d = \alpha^\phi, w = \beta^\phi, r = \gamma^\phi$$

This scaling algorithm enhances the efficiency of a model without necessarily requiring a high cost of computation.

The EfficientNetV2 is used in identifying the visible symptoms of cattle diseases using images in the proposed system.

The model processes cattle images and produces multi-class predictions using a softmax classifier:

$$P(y = i | X) = \frac{e^{z_i}}{\sum_j e^{z_j}}$$

Despite the complexity of deep learning architectures, EfficientNetV2 achieves improved accuracy and faster convergence compared to traditional CNN models. Its design makes it well suited for intelligent livestock health monitoring applications.

4 Results and analysis

4.1 Milk Quality Analysis Results

Multiple machine learning classifiers were evaluated for milk quality prediction using the structured dataset containing 10,001 samples and eight attributes

Table 2. Comparison of Machine Learning Models for Milk Quality Prediction

Model	Accuracy	Precision	Recall	F1-Score
Random Forest	95	0.96	0.95	0.955
Naïve Bayes	91	0.90	0.89	0.895
KNN	72	0.70	0.72	0.71
Support Vector Machines	58	0.55	0.58	0.565

These were precision, recall and F1-score in order to have a comprehensive assessment that goes beyond accuracy. Random Forest model exhibited balanced classification of the data, especially in the discrimination of similar classes like medium and high quality milk.

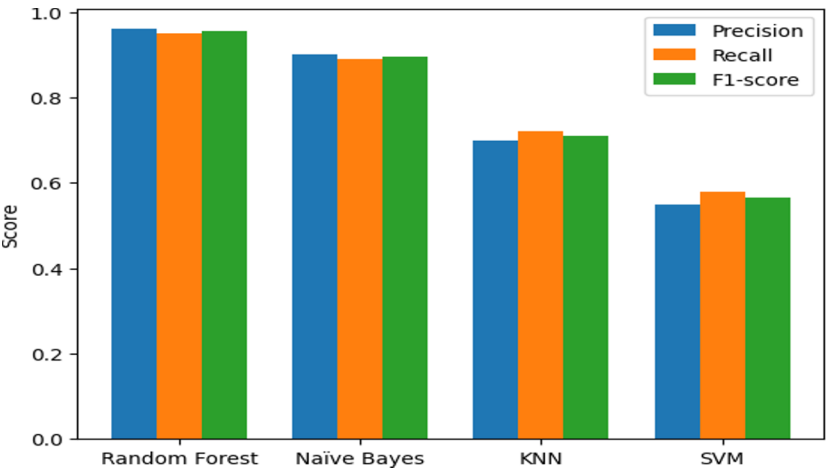


Fig. 3. Comparative Performance of Machine Learning Models

Figure 3 shows the relative analysis of machine learning models in terms of precision, recall, and F1-score. Random Forest model has always performed better than other models and this goes to confirm that it is a robust and reliable model to predict milk quality.

4.2 Cattle Disease Detection Results

For cattle health monitoring, deep learning models were evaluated on the image dataset containing three disease classes:

- Lumpy Skin Disease (LSD)
- Bovine Disease
- Foot and Mouth Disease (FMD)

Table 3. shows the comparison of Disease Detection Models.

Model	Accuracy	Precision	Recall	F1-Score
EfficientNetV2	95	0.96	0.95	0.955
MobileNetV2	88	0.89	0.87	0.88

EfficientNetV2 model performed better according to all the evaluation metrics. This is because of its efficient feature extraction and better classification accuracy through its optimized architecture and compound scaling method.

4.3 Visual Prediction Results

Figure 4, Figure 5 and Figure 6 shows diseases like Bovine, Lumpy, Foot and Mouth detected by the model respectively.

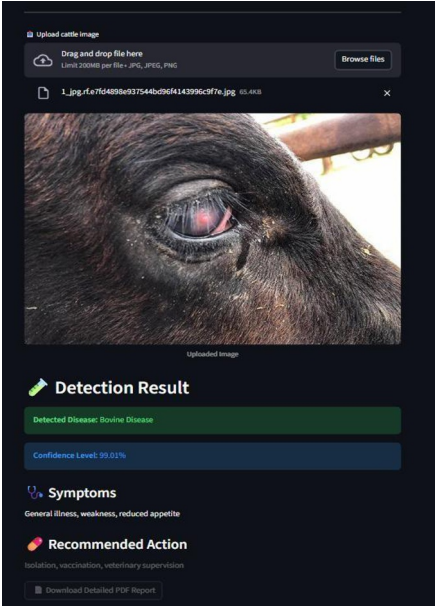


Fig.4. DETECTION OF BOVINE DISEASE

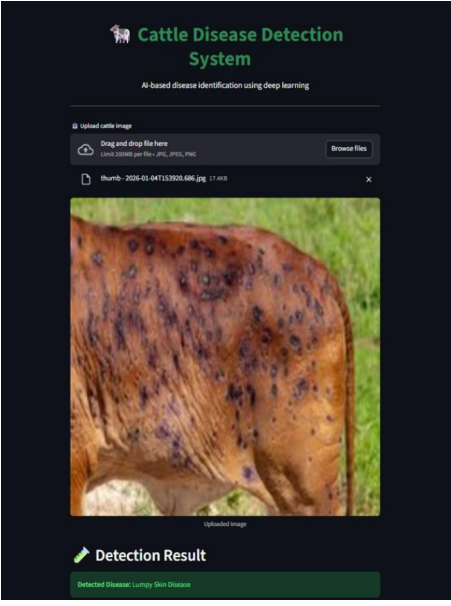


Fig.5. DETECTION OF LUMPY DISEASE

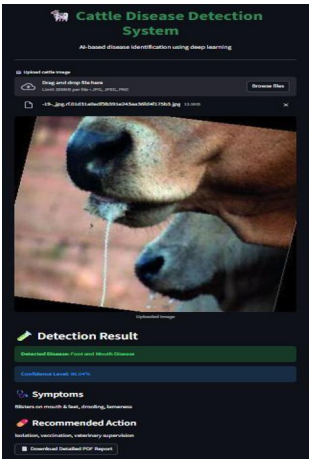


Fig. 6. Detection Food and Mouth

Milk Quality Prediction

pH Value

Temperature (°C)

Taste (0 / 1)

Odor (0 / 1)

Fat (0 / 1)

Turbidity (0 / 1)

Colour (0 / 1)

Select Model

Random Forest

Check Quality

Predicted Milk Quality: low

Fig. 7. Milk Quality Detected as Low

Milk Quality Prediction

pH Value

Temperature (°C)

Taste (0 / 1)

Odor (0 / 1)

Fat (0 / 1)

Turbidity (0 / 1)

Colour (0 / 1)

Select Model

Random Forest

Check Quality

Predicted Milk Quality: medium

Fig. 8. Milk Quality Detected as Medium

Milk Quality Prediction

pH Value

Temperature (°C)

Taste (0 / 1)

Odor (0 / 1)

Fat (0 / 1)

Turbidity (0 / 1)

Colour (0 / 1)

Select Model

Random Forest

Check Quality

Predicted Milk Quality: high

Fig. 9. Milk Quality Detected as High

Figure 7, Figure 8 and Figure 9 shows milk quality like Low, Medium and High detected by the model respectively

5 Conclusion

Shown to be very proficient in visual indicators. associated with common cattle illnesses, which makes them susceptible to automated detection. autonomous classification. For milk quality Comparison, relative examination of multiple machine learning. approaches also showed that the Random Forest algorithm is helpful. was better than its counterparts with a maximum accuracy of 95, indicating its successfulness in dealing with multi-variable, complex situations. datasets. Also, an advisory feeding module that is a chatbot. had been added to provide real time user friendly dietary. advice to the farm operators. Collectively, the proposed system increases automation of operations, increases predictive. performance, and allows scalable deployment, placing it as one of the viable and efficient solutions to modernize data-driven dairy farm management.

Declarations

Funding: The authors did not receive support from any organization for the submitted work.

Competing Interests: The authors have no relevant financial or non-financial interests to disclose.

Ethics Approval: Not applicable. This study does not involve human participants or animals.

Data Availability: The datasets analyzed during the current study are available from the corresponding author on reasonable request.

Author Contributions: All authors contributed to the study conception and design. Material preparation, data collection, and analysis were performed by Pratik Shingade, Adhik Sarak, Abhishek Walkoli, Lucky Ragde, Roshani Raut, and Radha Deoghare. All authors read and approved the final manuscript.

References

- [1] G. Krishna et al., “Exploring the Potential of Machine Learning for Early Cattle Disease Diagnosis,” in Proc. 5th Int. Conf. Inventive Research in Computing Applications (ICIRCA), IEEE, 2023, pp. 853–858.
- [2] Karthikeyan et al., “Lumpy Skin Disease Prediction Using Machine Learning,” in Proc. 2nd Int. Conf. Intelligent Data Communication Technologies, IEEE, 2024, pp. 887–892.
- [3] P. Bhardwaj et al., “Machine Learning Based Approaches for Livestock Symptoms and Diseases Prediction and Classification,” in Proc. Int. Conf. Communication, Computer Sciences and Engineering (IC3SE), IEEE, 2024, pp. 1035–1040.
- [4] R. U. Mhapsekar et al., “Hybrid Blended Deep Learning Approach for Milk Quality Analysis,” IEEE Trans. Emerg. Top. Comput. Intell., vol. 8, no. 3, pp. 2253–2268, 2024.
- [5] L.-F. Li et al., “Deep Learning-Based Skin Disease Classification Using Convolutional Neural Networks,” IEEE Access, vol. 8, pp. 123456–123467, 2020.

- [6] Y. Jia et al., “Behaviour-Based Classification of Dairy Cattle Health Using Machine Learning,” in Proc. Int. Conf. Sensing, Measurement & Data Analytics, IEEE, 2023, pp. 1–6.
- [7] J. Kamilaris and F. X. Prenafeta-Boldú, “Deep Learning in Agriculture: A Survey,” Comput. Electron. Agric., vol. 147, pp. 70–90, 2018.
- [8] S. Afshari-Safavi et al., “Deep Learning-Based Prediction of Lumpy Skin Disease in Cattle,” Comput. Electron. Agric., vol. 180, p. 105873, 2021.
- [9] S. Suparyati et al., “Livestock Disease Prediction Using Machine Learning Algorithms,” J. Anim. Sci. Technol., vol. 62, no. 4, pp. 512–521, 2020.
- [10] T. A. Putra et al., “Image-Based Cattle Disease Detection Using Convolutional Neural Networks,” Int. J. Adv. Comput. Sci. Appl., vol. 11, no. 9, pp. 240–247, 2020.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits any noncommercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

