YoloNet: Hybrid YOLO11 and EfficientNet for Lumpy Skin Disease Detection

Ansh Agrawal
School of Computer Science Engineering and Technology
Bennett University
Greater Noida, India

Neze Papreja
School of Computer Science Engineering and Technology
Bennett University
Greater Noida, India

Ayush Anand Singh
School of Computer Science Engineering and Technology
Bennett University
Greater Noida, India

Kashish Bhurani
School of Computer Science Engineering and Technology
Bennett University
Greater Noida, India

Sachin Minocha
School of Computer Science Engineering and Technology
Bennett University
Greater Noida, India

2025 3rd International Conference on Device Intelligence, Computing and Communication Technologies (DICCT) | 979-8-3315-4335-8/25/$31.00 ©2025 IEEE | DOI: 10.1109/DICCT64131.2025.10986576

***Abstract*— The Lumpy Skin Disease (LSD) is a very infectious and contagious disease which affects the lives of countless cattle each year. Machine/Deep Learning models like CNNs effectively detect this disease through images and track the progression of this disease in infected cattle. Various works have been done on this problem, using models such as MobileNetV2, ResNet50, Gradient Boost, Xception, Inception V3, and DenseNet121. This paper proposes a hybrid model, YoloNet, to detect the LSD virus in Cattle, using YOLO11 and EfficientNet-B0. The YoloNet model takes a pre-processed image and performs feature extraction using YOLO11 to generate a tensor. The generated tensor is given as input to EfficientNet-B0, which converts it to a feature map and classifies the image. The analysis of the YoloNet has been done on a dataset with 5 different classes of cattle, with 8014 total images. The accuracy of the YoloNet model, being 97.97%, is significantly better than other pre-existing state-of-the-art models.**

Keywords—Lumpy Skin Disease, YOLO, Yolo11, EfficientNet, Deep learning

I. INTRODUCTION

Livestock plays a crucial role in agriculture and serves as the foundation of the economy for several countries like India. It significantly impacts the country's gross domestic product and agricultural sector. In India, approximately 35.94% of all livestock are cattle [1]. Cattle and livestock not only supply the country with food, such as dairy products but also create job opportunities and a reliable source of income for isolated urban areas. The mishandling and inefficient management of livestock can lead to their scarcity, leading to a severe shortage of nourishment and hindering livelihood. The awareness related to cattle health is extremely low, which results in thousands of cattle deaths every year. In 2022, over 155,366 cattle were impacted or died due to the Lumpy Skin Disease (LSD) [2].

The Lumpy Skin Disease (LSD) initially originated in Zambia, a country in Africa and is a extremely infectious and dangerous disease [3]. It is triggered by the Lumpy Skin Disease Virus (LSDV), belonging to the Capripoxvirus genus and it majorly affects cows and water buffalos [4]. The Lumpy Skin Disease Virus is extremely contagious and can survive for prolonged periods of time since it is adjustable in various environments, making it extremely difficult to eradicate if it remains undetected in the early stages [5]. LSD can critically affect cattle in various ways. It can lead to infertility in the cattle, lead to abortions, and even harshly decrease calf survival rates [6]. Further, it can also result in reduced milk production, leading to severely consequential economic losses for the farmers. This disease can also make the indisposed cattle more vulnerable to other diseases like Foot and Mouth disease and cattle tuberculosis.

Once the disease has reached its peak, there is no particular cure, but early detection can help control it through proper treatment. The symptoms of LSD include high fever, skin lesions, and nodules on the body. Camera pictures and video clips can identify symptoms like raised nodules, watery eyes, scabs in the skin, and lesions on oral and nasal cavities [7].

Machine learning proves to be an effective and non-invasive method for the early detection of LSD [8]. The use of convolutional neural networks (CNN) further improves the detection. These models can find subtle changes in images and can thus detect underlying LSD symptoms in the early stages [9]. Such early detection can prevent the transmission of this disease and could even help to cure it. Image analysis can also help track the progression of this disease in cattle.

Various authors have worked on the early detection of LSD. Sivamurugan and Uthayan [10] compared different models, including Xception, InceptionV3, VGG19, DenseNet121, ResNet50, and MobileNetV2, on a dataset of 6000 images divided into three categories. MobileNetV2 model has shown the highest accuracy of 0.9182.

Karthikevan et al. [11] compared ensemble-learning methods such as random forest, extra trees classifier, sequential CNN, and pre-trained models (Dense121, Resnet50) on a dataset of 700 images for the classification of healthy and LSD cattle to get the highest accuracy of 96.85% using Resnet50.

Ujjwal et al. [12] have applied gradient boosting, KNN, decision tree, random forest, SVM, naive Bayes, Adaboost, and CNN for lumpy skin disease occurrence detection. They have analyzed the mentioned techniques on a dataset consisting of 18603 instances and 16 features. The analysis presented the better performance of the random forest algorithm as compared to other mentioned algorithms.

Saqib et al. [13] used the MobileRMSNET model and the RMSprop optimizer. Their hybrid model shows an accuracy of 95% on a dataset consisting of 464 images of healthy cows and 329 images of infected cows. Saha [9] has also compared different models, including CNN, DenseNet, MobileNetV2, Xception, and InceptionResNetV2 to categorize the images, wherein MobileNetV2 has been able to achieve an accuracy of 96% and an AUC score of 98%.

The analysis done by the various authors has shown significant performance by the various individual models. However, hybrid deep learning models to identify the Lumpy Skin disease efficiently is still an open area for research. This work hybridizes two models, i.e., YOLO11 and EffecientNet. This hybridization is done due to complimentary attributes of two models, i.e., object detection and classification, respectively. The contributions of this paper are:

- Design a hybrid model using pre-existing models namely, YOLO11 and EffecientNetB_0.
- Classify the cattle into healthy and those affected by the Lumpy Skin disease using the proposed hybrid model.
- ·An in-depth comparative analysis of different pre-existing models to prove the efficiency and the usefulness of the proposed model.

The remaining paper is divided into four sections. Section II discusses the model architecture of YOLO11 and EfficientNet. The following section explains the proposed model in detail. Section IV represents the results and discussions including the comparative analysis of various models. Finally, Section V concludes the findings of the paper.

II. Preliminary Studies

The proposed solution is based on two different models, i.e., YOLO11 and EfficientNet. This entire section focuses on the architecture and workings of both these models.

A. *YOLO11*

YOLO11 is an extended and improved version of YOLO (You Only Look Once), which is a region-based Convolutional Neural Network (CNN), [14] developed by Ultralytics. It aims at providing better speed, precision, and accuracy in terms of advanced real time object detection. Various versions of the YOLO have been used for the early and timely detection of the lumpy skin disease in cattle. YOLO11 exhibits a highly efficient architecture, as shown in Fig. 1, with C3k2 (Cross stage spatial with kernel size 2) blocks, Spatial Pyramid Pooling Fast (SPPF), and cutting-edge attention mechanisms like C2PSA. This efficient architecture has the capability to process spatial information while maintaining its high-speed inference [15].

YOLO11 presents a remarkable enhancement by replacing the C2f (Coarse-to-fine) block with the C3k2 (Cross stage spatial with kernel size 2) block. The C3k2 block is designed to be quicker and more effective, improving the overall performance of the feature accumulation process [16].

The C2PSA (Convolutional block with Parallel spatial attention) block improves the spatial attention in the feature maps, enhancing the model's focus on the key areas of the image. This model focuses on important regions of interest by pooling features together [17].

B. *EfficientNet*

EfficientNet is a really powerful convolutional neural networks which strikes a balance between accuracy and efficiency [18]. EfficientNet balances network width, depth, and resolution uniformly using a set of fixed grading coefficients and improves the level of efficiency without compromising on accuracy [19]. EfficientNet efficiently reduces the use of computational resources in order to develop a new machine learning model.

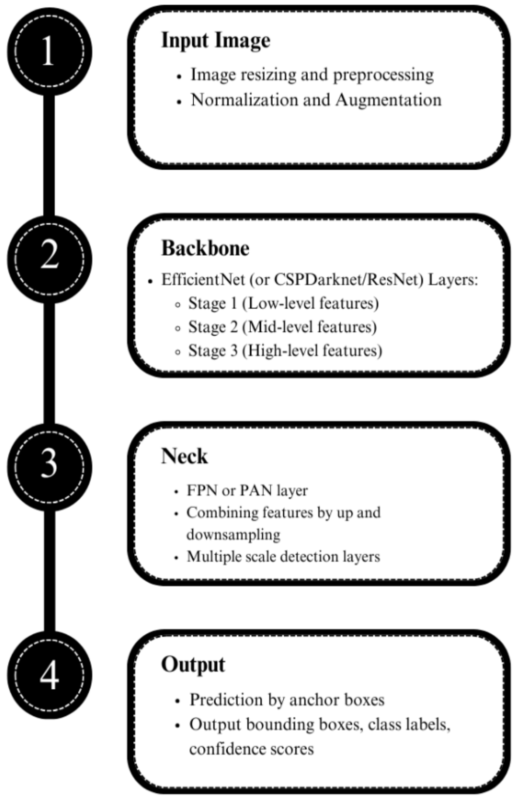

Fig. 1. YOLO11 Skeleton

The architecture of EfficientNet shown in Fig. 2 is constructed on mobile inverted bottleneck convolution (MBConv). It is a mixture of depth-wise distinguishable convolutional layers and inverted outstanding blocks which is similar to MobileNetV2. It emphases on reducing computational cost while maintaining accuracy [20]. EfficientNet uses a linear activation function instead of ReLU which helps retain information that would have been lost due to ReLU's nonlinear function.

Furthermore, EfficientNet comprises of the SE (Squeeze and Excitation) block, which helps the model focus on key characteristics while suppressing the other features [21]. The SE(Squeeze and Excitation) block uses global average pooling to compress the feature map's spatial dimensions to a single channel. The average layer is then followed by two fully linked layers that help to standardize the original feature maps, leading to improved representational power.

The EffecientNet model ranges from B0 to B7 [22]. The increase in complexity from B0 to B7 provides a trade-off between accuracy and computational cost. The model from B0 to B7 can be selected according to the complexity of various applications, including medical imaging, semantic segmentation, and object detection.

III. Proposed Work

This work uses the YOLO11 and the EfficientNet-B0 to design a hybrid architecture named YoloNet that provides better detection of lumpy skin disease. YOLO and

EfficientNet were chosen because of their high object detection and classification features.

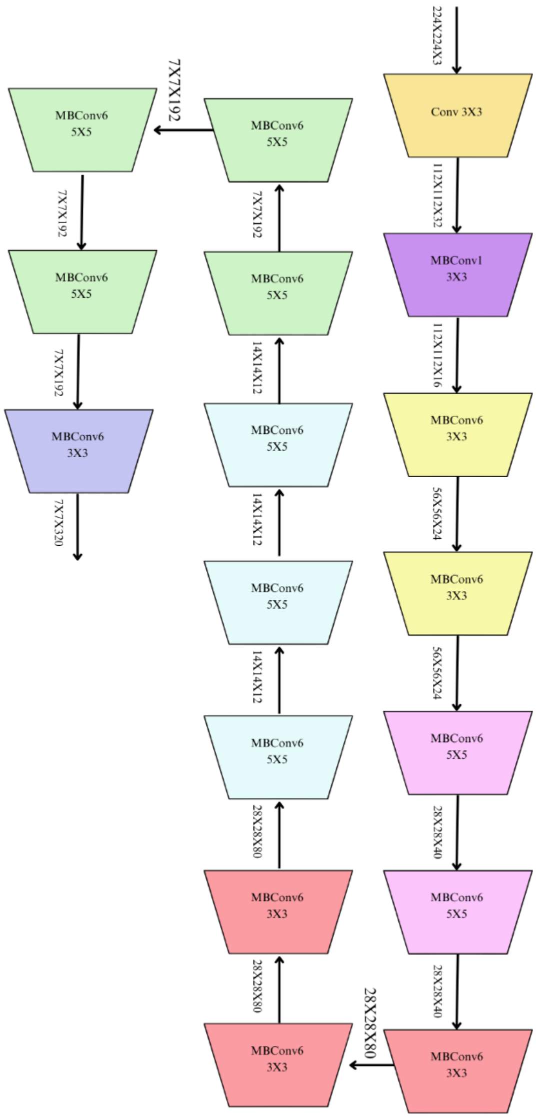

Fig. 2. EfficientNet Architecture

The proposed YoloNet works in the following phases, as shown in Fig. 3. The first phase processes the input image, followed by phase 2, which uses the YOLO11 to detect the region of interest. Phase 3 gets the phase 2 output as input and classifies the cattle as healthy or diseased using EfficientNet-B0. The detailed workings of the proposed YoloNet are as follows.

Phase 1

This phase ensures consistency and compatibility with the model's input requirements. In this phase, an input image is pre-processed so that it can be given as input to YOLO11 in phase 2. The input image to phase 1 is resized into 224 x 224 pixels and then converted to tensors. After that, the final batch preparation using all the processed images and mapped labels is done.

Phase 2

The output of phase 1, i.e., the processed image, is given as input to this phase. This input is processed using YOLO11. Input first goes through the Yolo backbone, which helps in feature extraction, and then to the Yolo neck, which helps in mapping features by combining different scales. Finally, through Yolo's head, which helps in predicting bounding boxes at the end.

Phase 3

The output from YOLO11 is then fed as the input to EfficientNetB0. The input initially goes through the EffecientNetB0 backbone which helps in the feature extraction, then EffecientNetB0 helps capture essential features by making feature maps. After this, it goes to a flattened layer, which helps convert multi-dimensional feature maps to one-dimensional vectors for the fully connected layers. The fully linked dense layer is used with 256 units. This layer uses the ReLU activation function to present non-linearity and learn complex patterns from the flattened feature vector. This layer converts the high-dimensional data into a smaller size while preserving key information.

The output from the fully connected layer is then used as input into the classification output layer. This layer introduces ‘n’ class softmax scores, which then lead to the classification of the data into each of the ‘n’ classes. Based on the softmax scores, the model then makes a prediction for the specific class label, classifying the lumpy skin disease. The results of the proposed model and corresponding analysis are explained in the next section.

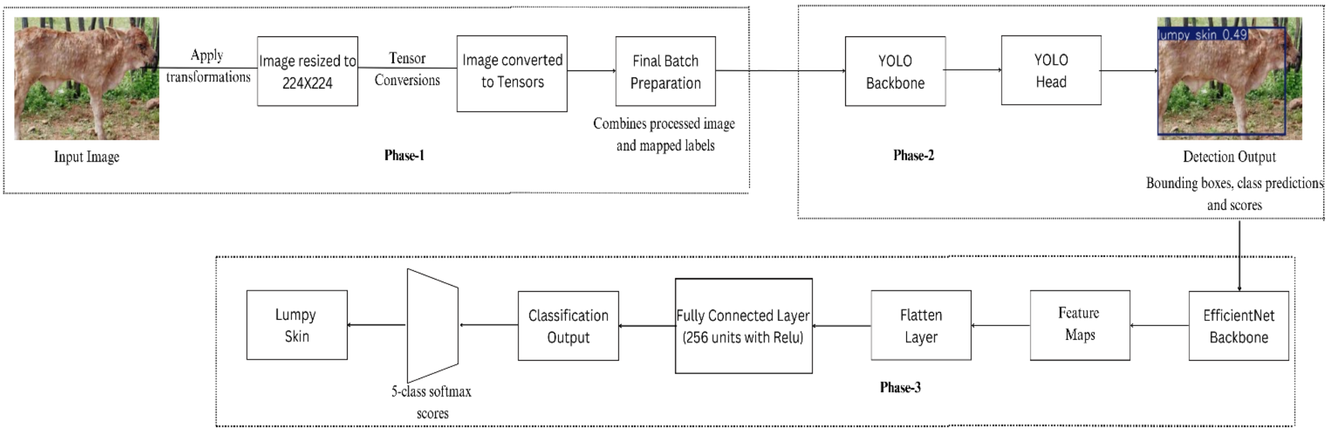

Fig. 3. Proposed YoloNet Architecture

IV. Results and Discussions

This section presents the details of the implementation and performance of the proposed hybrid model, YoloNet. It also explains the performance metrics and compares them with baseline models and advanced models. The training and testing processes were done on a system with an Intel i7-12700H, 16 GB RAM, NVIDIA RTX 3060, and 6 GB VRAM configuration.

The open-source dataset [23] used for this work has five classes, i.e., infected foot, mouth disease infected, normal healthy cow, normal mouth, and lumpy skin disease. It has 8014 images that are split into an 88-6-6 ratio for training, testing, and validation, respectively. The dataset images vary in size, so they are resized into 224 X 224 pixels and later converted to tensor arrays for smooth execution across the models.

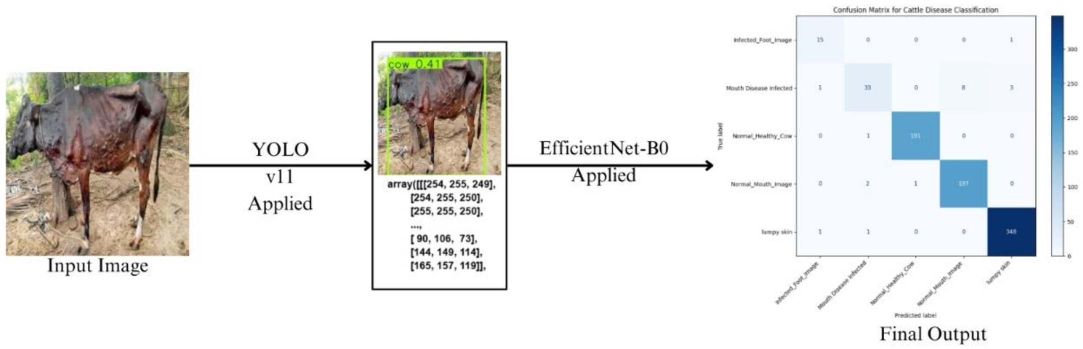

Fig. 4. Implementation of YoloNet

In the proposed YoloNet model shown in Fig 4, the preprocessed image is given as input to Yolo11. Yolo11 finds the ROI by marking the bounding boxes and generating corresponding tensors, as presented in Fig 4. This tensor is passed as input to EfficientnetB_0, which performs feature extraction using convolutional layers and the fully connected layers to classify the image into one of the five classes, namely Infected Foot Image, Mouth Disease Infected, Normal Healthy Cow, Normal Mouth Image, Lumpy Skin Disease. The performance of YoloNet is analyzed using the Accuracy, F1-score, Precision, and Kappa coefficient shown in Table 1.

TABLE I. Performance comparison of YoloNet

Models	Accuracy	Precision	KAPPA	F1-score
MobileNetV2	96.85%	0.86	0.87	0.88
DenseNet121	85%	0.72	0.76	0.76
ResNet50	95%	0.96	-	0.94
InceptionV3	91.3%	0.82	0.82	0.85
Gradient Boost	95.7%	0.95	-	0.95
YOLO11n	80.02%	0.69	0.92	0.87
EfficientNet-B0	95.02%	0.91	0.92	0.87
Proposed Model (YoloNet)	97.7%	0.95	0.96	0.95

Table 1 signifies that the YoloNet model performs better than baseline models, i.e., Yolo11 and EfficientNet. YoloNet reflects an improvement of 2.68% and 17.68% in accuracy as compared to Yolo11 and EfficientNet, respectively. YoloNet shows improvement in other parameters also.

YoloNet model performance is also compared with five models, MobileNetV2, DenseNet121, ResNet50, InceptionV3, and Gradient Boost. The comparison is also shown using Fig. 5, 6, and 7 for parameters accuracy, precision, and F1-score.

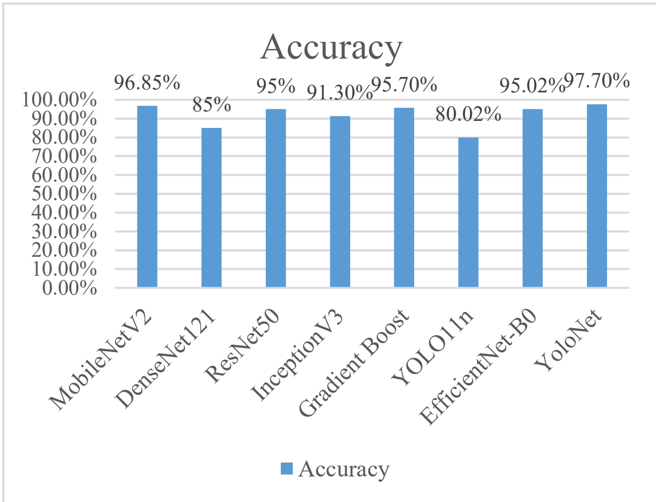

Fig. 5. Performance Comparison in terms of Accuracy

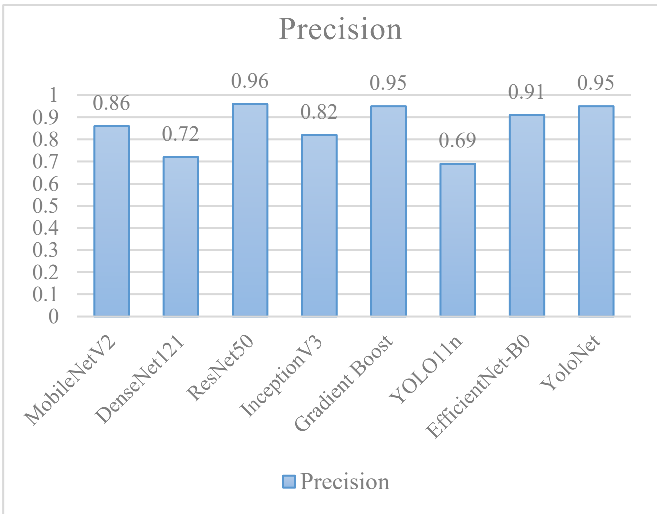

Fig. 6. Performance Comparison in terms of Precision

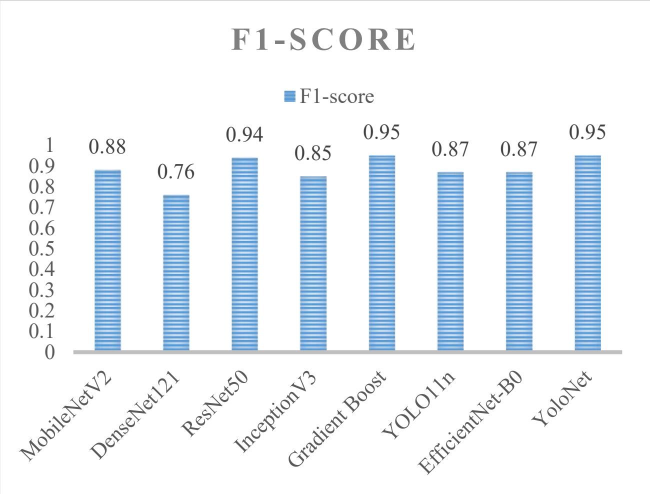

Fig. 7. Performance Comparison in terms of F1-score

Fig. 5, 6, and 7 compare the accuracy, precision, and F1-score, respectively. YoloNet exhibits the best accuracy and F1-score, while the precision of ResNet50 is higher than the YoloNet. However, higher values of F1-score and accuracy displays better performance of the proposed model than all the other models. This better performance is due to the ensembling of two complementary models, which results in better object detection and classification.

V. Conclusion

This work proposes a hybrid model named YoloNet using Yolov11 and EffecientNet for lumpy disease detection in cattle. YoloNet combines the Yolov11 object detection features and EffecientNet classification properties to generate an efficient model. This model achieves an accuracy of 97.7%, which exhibits an improvement of 2.68% and 17.68% as compared to EfficientNet and Yolo11, respectively, over the same dataset. Furthermore, this work shows a performance comparison of five models, MobileNetV2, DenseNet121, ResNet50, InceptionV3, and Gradient Boost, with the proposed model using accuracy, F1-score, precision, and Kappa as analysis parameters. The improved performance of the proposed model than all the pre-existing models available shows the significance of the work. In the future, this model can be optimized to deploy on low-power devices like mobile phones for real-time detection in rural areas. It can be extended to analyze video clips.

References

[1] A. K. Vinod, "An Economic Analysis of Indian Livestock Sector using livestock census 2019," 2022.

[2] Open Government Data (OGD) Platform India, "State/UTs-wise Details of the Death of Cattle due to Lumpy Skin Disease (LSD) in the Country during 2022," Answers Data of Rajya Sabha Questions for Session 258.

[3] L. Khan & M. Gul, "A comprehensive review on lumpy skin disease," Pure and Applied Biology, vol. 13, no. 3, Sep. 2024, doi: 10.19045/bspab.2024.130031.

[4] G. Moudgil, J. Chadha, L. Khullar, S. Chhibber & K. Harjai, "Lumpy skin disease: Insights into current status and geographical expansion of a transboundary viral disease," Microb Pathog, vol. 186, p. 106485, Jan. 2024, doi: 10.1016/j.micpath.2023.106485.

[5] M. Akther, S. H. Akter & S. Sarker, "Global Burden of Lumpy Skin Disease, Outbreaks, and Future Challenges," Viruses, vol. 15, no. 9, p. 1861, Aug. 2023, doi: 10.3390/v15091861.

[6] Y. R. Khan, A. Ali & K. Hussain, "A review: Surveillance of lumpy skin disease (LSD) a growing problem in Asia," Microb Pathog, vol. 158, p. 105050, Sep. 2021, doi: 10.1016/j.micpath.2021.105050.

[7] E. Mulatu & A. Feyisa, "Review: Lumpy Skin Disease," J Vet Sci Technol, vol. 09, no. 03, 2018, doi: 10.4172/2157-7579.1000535.

[8] N. Niloy & T. Bhowmik, "Exploring Machine Learning Techniques for Symptom-Based Detection of Livestock Diseases," May 2024.

[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, p. e34242, Jul. 2024, doi: 10.1016/j.heliyon.2024.e34242.

[10] V. Sivamurugan, K. R. Uthayan, & V. Thanikachalam, "Cow Disease (LSD) Classification System for predicting different Severity levels.," International Journal of Early Childhood Special Education, vol. 14, no. 8, 2022.

[11] S. R. Karthikevan, Madhanakumar & V. D. Mekala, "Lumpy Skin Disease Prediction Using Machine Learning," in 2024 2nd International Conference on Intelligent Data Communication Technologies and Internet of Things (IDCIoT), IEEE, Jan. 2024, pp. 887–895. doi: 10.1109/IDCIoT59759.2024.10467939.

[12] N. Ujjwal, A. Singh, A. K. Jain & R. G. Tiwari, "Exploiting Machine Learning for Lumpy Skin Disease Occurrence Detection," in 2022 10th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO), IEEE, Oct. 2022, pp. 1–6. doi: 10.1109/ICRITO56286.2022.9964656

[13] T. B. O. Muhammad, S. S. Muhammad, & I. Muhammad, "Lumpy skin disease diagnosis in cattle: A deep learning approach optimized with RMSProp and MobileNetV2," PLoS One, vol. 19, no. 8, p. e0302862, Aug. 2024, doi: 10.1371/journal.pone.0302862.

[14] N. Jegham, C. Y. Koh, M. Abdelatti & A. Hendawi, "Evaluating the Evolution of YOLO (You Only Look Once) Models: A Comprehensive Benchmark Study of YOLO11 and Its Predecessors," Oct. 2024.

[15] R. Khanam & M. Hussain, "YOLOv11: An Overview of the Key Architectural Enhancements," Oct. 2024.

[16] M. A. Alif, "YOLOv11 for Vehicle Detection: Advancements, Performance, and Applications in Intelligent Transportation Systems," Oct. 2024.

[17] R. Khanam, and M. Hussain, "Yolov11: An overview of the key architectural enhancements," arXiv preprint arXiv:2410.17725, 2024.

[18] B. King & M. Rizkalla, "Enhanced Multiple Dense Layer EfficientNet," Purdue University, 2024

[19] M. M. Islam, Md. A. Talukder, M. A. Uddin, A. Akhter & M. Khalid, "BrainNet: Precision Brain Tumor Classification with Optimized EfficientNet Architecture," International Journal of Intelligent Systems, vol. 2024, no. 1, Jan. 2024, doi: 10.1155/2024/3583612

[20] V.-T. Hoang, & K.-H. Jo,"Practical Analysis on Architecture of EfficientNet," in 2021 14th International Conference on Human System Interaction (HSI), IEEE, Jul. 2021, pp. 1–4. doi: 10.1109/HSI52170.2021.9538782.

[21] S. Ucan, M. Ucan & M. Kaya, "Deep Learning Based Approach with EfficientNet and SE Block Attention Mechanism for Multiclass Alzheimer's Disease Detection," in 2023 4th International Conference on Data Analytics for Business and Industry (ICDABI), IEEE, Oct. 2023, pp. 285–289. doi: 10.1109/ICDABI60145.2023.10629380.

[22] M. Pathak, D. Chaudhary, K. Sharma, A. K. Sharma, A. Gupta, & B. K. Sharma, "A Robust EfficientNet Architecture for Brain Tumor Classification and Identification Using MRI Image," in 2023 11th International Conference on Intelligent Systems and Embedded Design (ISED), IEEE, Dec. 2023, pp. 1–5. doi: 10.1109/ISED59382.2023.10444597.

[23] M. Rahaman, Cattle_disease Detection Dataset. Roboflow Universe,2023.[Online].Available:https://universe.roboflow.com/mdzillur-rahaman-rohan/cattle_disease-detection[Accessed: 2024, Nov. 9].