Deep Learning based Real-Time Animal Detection using MobileNetV2

D.Manikandan[1*], M. Swarna[2], D. Nivetha [2], Dharini Kavya [2]
[1] Assistant Professor, [2] UG Scholar, Department of Computer Science and Engineering,
School of Engineering,
Vels Institute of Science, Technology and Advanced Studies
Corresponding Email: pugalanthimanikandan40@gmail.com

ABSTRACT: In this modern world, Artificial intelligence has emerged as a powerful tool in wildlife conservation, agricultural land protection, and environmental monitoring. The study proposes a real-time animal detection and classification system by utilizing the MobileNetV2 deep learning model. MobileNetV2 is selected for its lightweight architecture, high computational efficiency, and superior accuracy, making it suitable for real-time applications. The model achieves an outstanding accuracy of **99.62%**, which outperforms with other deep learning models. The proposed system enhances real-time monitoring by detecting and classifying animals in various environments, aiding in wildlife protection, mitigating human-animal conflicts, and securing agricultural lands. By providing instant alerts, the system enables rapid decision-making during emergency situations, ensuring safety and effective intervention. The system is designed to be scalable, cost-effective, and adaptable to different terrains, making it an ideal solution for large-scale deployment in conservation areas and farmlands.

Aim: AI-driven approach promotes proactive environmental monitoring, helping authorities track endangered species, prevent poaching, and maintain ecological balance. The integration of deep learning in animal detection not only improves efficiency but also revolutionizes traditional monitoring methods, ensuring a sustainable coexistence between human activities and wildlife. This research highlights the potential of AI in preserving biodiversity and safeguarding natural ecosystems.

Keywords: MobileNetV2, Deep learning, Animal detection, Fine-tuned, Decision making

Background:

MobileNetV2 is an advanced CNN architecture focused on accomplishing effective image classification, object recognition, and performing other computer vision tasks in devices with limited resources. It is developed as a successive improvement of MobileNet which has been built with special emphasis on the real time applications. It is useful in applications where there are limits with regard to power and computational resources such as mobile phones, IoT devices, and edge computing systems. In addition, MobileNetV2 utilizes an inverted residual structure that possesses linear bottlenecks which aids the network to learn features from data with low computation cost. Such design helps in achieving both in efficiency and accuracy and suitable for applications that involve detecting animals in natural settings.

MobileNetV2 is a recent achievement targeting animal detection tasks that has become popular. It exhibits a relatively better performance with respect to accuracy along with the reduced computational complexity in real time scenarios [1]. MobileNet-SSD V2 with TensorFlow Lite was proved to be effective in identifying many species even under harsh conditions of low illumination and complicated situations and delivering high identification performance [2]. MobileNetV2 coupled with SSD was able to achieve detection rates of 80% for tigers 89.47% for jaguars and 92.56% for elephants at the rate of 2-3 frames in a second, demonstrating its efficiency for deployment at massive levels [3]. MobileNetV2's architecture and high efficiency has suited it for deployment in resources constrained environments such as mobile applications. As reported, with the help of transfer learning and appropriate model modification, MobileNetV2 performs well on tasks including the classification of fruits. [4]. An integrated hybrid approach based on MobileNetV2 and Vision Transformer performed well on extracting the features and managed to attain an accuracy of 96.94% in sheep face recognition while targeting a lightweight structure suitable for edge devices [5].

MobileNet-SSD V2 was also noted for its usability in changing environments which translated into improved species recognition accuracies on several datasets [6]. These findings confirm MobileNetV2 as a relevant model in most tasks since it can improve the results of most applications without redundancy, such as real-time animal detection and monitoring. The Deep convolution neural networks (DCNNs) were used to identify animals with high accuracy by learning features from training images and predicting efficiently. This is with regard to streamlined efforts for wildlife monitoring and ecosystem conservation [7]. Convolution Neural Network (CNN) based models also prevent animal-vehicle collisions with minimal false positives, providing real-time alerts as well as safe passage for both man and animals [8]. Although these works focus attention on the use of CNN architectures, they provide the base to continue discovering more advanced lightweight networks such as MobileNetV2 in real-time animal detection systems. [9]. An animal incursion detection system uses Faster RCNN and SSD-MobileNet models the former proved to be more efficient and accurate. Training of models was done using TensorFlow; the efficiency of the system assures real-time detection of animals and the prevention of incursions. Therefore, the results established sound ability on labeled datasets, indicating real use in surveillance and wildlife management. [10].

MobileNetV2 model reduces the computation time and improved accuracy using knowledge learned from large datasets like Image Net. MobileNetV2 fills the gap in currently used models that mostly train from scratch or use simpler architectures and captures complex features where intensive data augmentation techniques are used for better generalization for filling some part between models. Fine tuning of first 100 layers with MobileNetV2 makes the model for specific and accurate when compared with the previous approaches. The hyper-parameter tuning includes the batch size of 32 and smaller learning rate of 0.0001 which leads the better convergence without any updates in weight which are often neglected area of study.

Experimental

The Proposed animal detection model is used to identify and detect the specific animal using Mobile NetV2 model in which it helps at the time of emergency situations. The animal data samples are collected and pre-processed where the feature is extracted and then model is training using multiple deep learning model. The proposed model is divided into three modules such as data collection, data preprocessing, feature extraction and model training.

Data Collection: The dataset comprises of different animal images taken in different geographical region, i.e. the images are with the natural settings like farms and zoos. The dataset also contains a variety of weather condition such as rainy, sunny and cloudy with different time periods such as morning, midday and evening. The dataset is organized with multiple classes in which each class comprises of multiple images are shown in the figure 1. Forty-Eight classes are considered with the 60 images each, where both wild animals and domestic animals are considered. A synthetic dataset is created and classes are classified according to the animals.

Data Preprocessing: The process begins with Data Acquisition, where data is collected from various sources such as online repositories, cameras, and sensors. This data is then categorized to ensure it is properly structured for subsequent processing. Following acquisition, Image Resizing is performed, where the images are downscaled to a standard size of 244x244 pixels. Padding is applied to maintain the original aspect ratio of the image, which helps to normalize the data and ensures consistency across all input images. To reduce noise, unnecessary parts of the images are cropped or removed. Additionally, color space conversion is carried out to transform the images into grayscale, reducing the complexity of the data and focusing on the most relevant features for accurate identification.

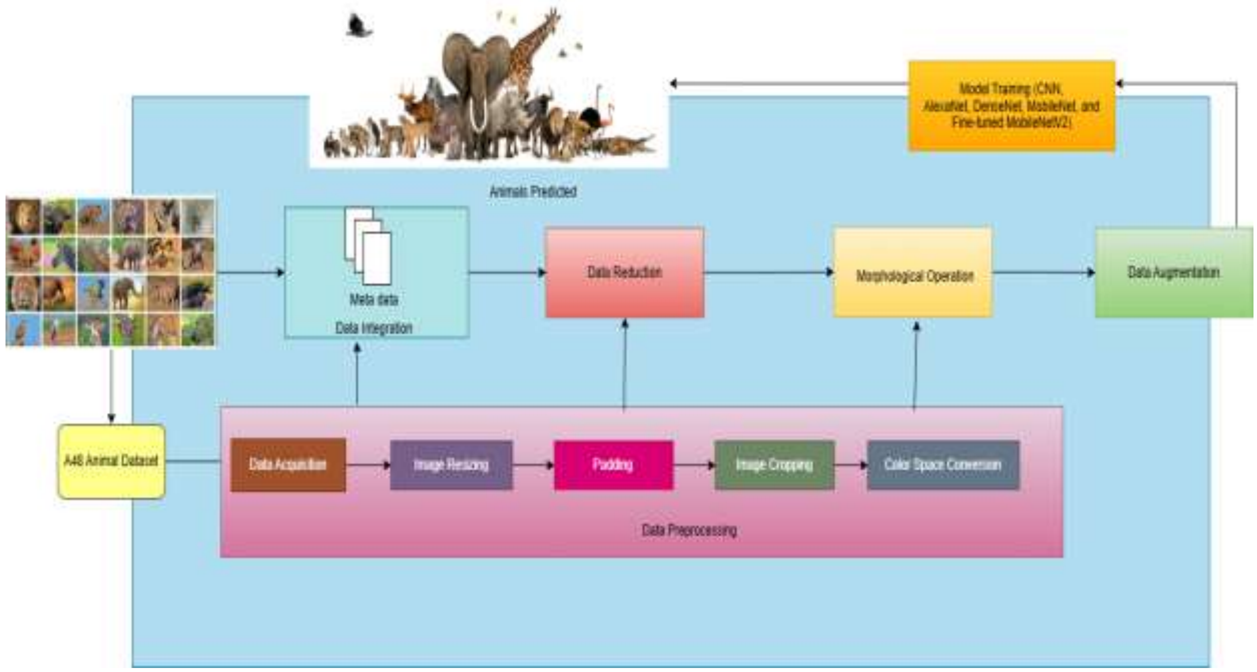


Figure 1. Proposed Deep Learning based Real-Time Animal Detection using MobileNetV2

Data cleaning follows, during which any unwanted noise is eliminated, ensuring the images are of higher quality. After cleaning, Data integration takes place, where meta-features (e.g., species attributes) are integrated into a single and cohesive dataset. Next, Data reduction is applied to remove any improper, redundant, or irrelevant data, and Morphological operations are performed to refine the images, removing unnecessary information and improving the clarity of the relevant features.

To further enhance the dataset, Data augmentation is employed. This technique helps to increase the dataset size and improves the model's robustness by introducing variations such as noise reduction, scaling, and rotations. Augmentation also helps to prevent overfitting by artificially expanding the diversity of the dataset. Additionally, oversampling is performed to balance the class distribution, ensuring the model doesn't become biased toward over represented classes. This entire process, from acquisition to augmentation ensures the data is cleaned, diverse, and ready for optimal model performance.

Model Training After preprocessing, the data undergoes classification and detection. Four deep learning models are utilized for training: Convolutional Neural Network (CNN), Alex Net, Dense Net, Mobile Net, and MobileNetV2. The convolutional layers serve as feature extractors, with max pooling layers reducing dimensionality and fully connected layers classifying the animal species accurately. Alex Net enhances the traditional CNN, incorporating an input layer for images of size 227×227times 227×227 with three convolutional layers. Dense Net employs multiple dense connections to ensure optimal feature flow. Images are resized to 224×224 times 224×224, and the model leverages the pre-trained DenseNet-169 as a feature extractor, fine-tuned with the dataset. Mobile Net, a lightweight convolutional neural network, is specifically designed for embedded and mobile devices. It follows the same procedure to classify animal images. MobileNetV2, an improved version of Mobile Net, excels in both classification and detection, offering better accuracy compared to other models.

Results and Discussion

The synthetic A48 animal dataset is extracted from the online repositories and some images are collected as a real-time image of animals. Antelope, Bear, Bison, Boar, and Cat, chimpanzee, cow, coyote, deer and other 39 animal images are collected, where each animal class contains 60 different images and labelled with the specific class. Preprocessing is performed to reduce the noise in the image and features are extracted using the convolutional layer. TensorFlow is used to develop the deep learning model where keras is used as API for the rapid development. The dataset is divided into train set and test set, where 80:20 ratio is considered. Adams optimizer is used to compute the adaptive learning rates of each layer. CNN and pre-trained model of AlexNet, DenseNet, MobileNet and MobileNet V2 are used to train the model in which it classifies the images of the animals with the accurate species, respectively. CNN fails in producing the better accuracy when compared with the other deep learning models.

Different performance metrics such as precision, recall, accuracy and F1 score are compared. CNN provides the accuracy of 0.87, whereas AlexNet provides 0.84, DenseNet provides 0.89, MobileNet produce 0.95 as accuracy and it is better than other 3 models. MobileNetV2 produce the 96% accuracy in which it highlights a slight difference between MobileNet. To get better accuracy, MobileNetV2 is fine-tuned, which improves an accuracy of 98% and reduces the loss of 8% as shown in the figure 2 and it is far better than other models considered.

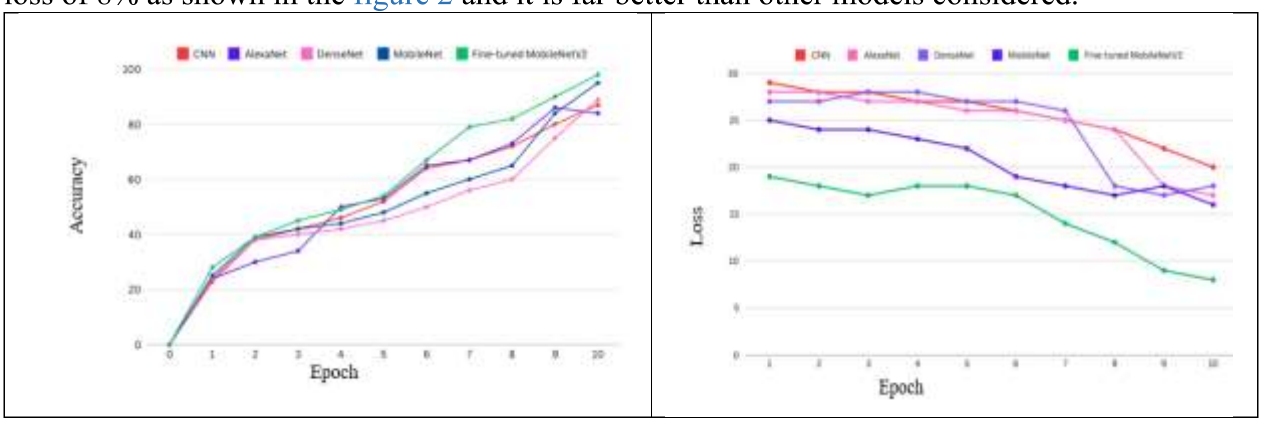
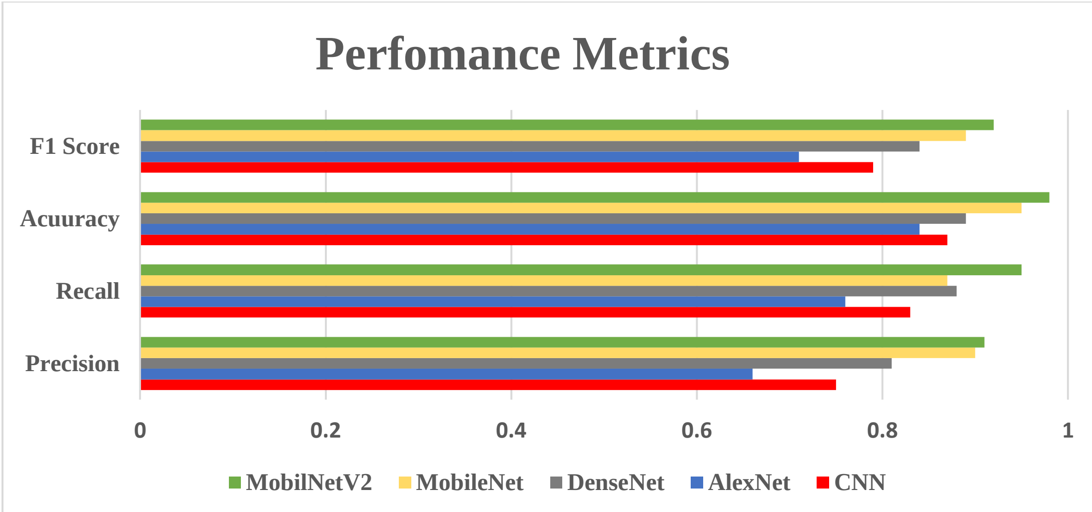


Figure 2. Accuracy and Loss Comparison

Other performance metrics such as precision, recall and F1 score are identified as shown in the figure 3. Precision, recall and F1 score of MobileNetV2 is 0.91 whereas other models fail to get the better results.



Model	Precision	Recall	Accuracy	F1 Score
CNN	0.75	0.83	0.87	0.79
AlexNet	0.66	0.76	0.84	0.71
DenseNet	0.81	0.88	0.89	0.84
MobileNet	0.90	0.87	0.95	0.89
MobileNetV2	0.91	0.95	0.98	0.92

Figure 3. Performance Metrics of the Models

MobileNetV2 based animal detection model experimental research shows that it provides the high accuracy for the real-time animal detection. The models are trained on an animal image dataset and used data augmentation techniques like rotation, width/height shifts and horizontal flip to increase the variability of the training data. This helps the model to generalize better and reduces the overfitting of the data. MobileNetV2 reduces the computational time and improves the accuracy when the model starts with knowledge from a large dataset like ImageNet. MobileNetV2 is fine-tuned by setting the 100 layers free which allows the model to learn specific features of the animals. The model is trained with the 10 epochs with frozen and fine-tuned layers and it helps to converge to identify the specific animal with 98.67% training accuracy and 86.70% validation accuracy, where the model does an incredibly well on training data and testing data stand with 97.8% accuracy. Additionally, a batch size of 32 with smaller learning rates during the model helps in smooth convergence without drastic changes in weights, which ensures in stable trainings and performances of models.

An elephant image is evaluated and classified with 98% accuracy and it improves the reliability of the model as shown in the figure 3. It shows that MobileNetV2 model can make the prediction with the complex diverse images with minimal error. Additionally, the prediction is made with the high confidence where this enhances the reliable of the image.

Conclusion

The MobileNetV2 model is quite efficient and effective to achieve real time detection of animals with significant reduction in cost and resources. ImageNet is used as the source to obtain the pre-trained weight using transfer learning. MobileNetV2 model provides better accuracy by identifying the accurate animal with species where data augment techniques help more to identify the specific features of the animal. Many directions can be made in prospect of detecting the animal in real time at emergency situation. The dataset can be increased and include a wide variety of species with different environmental condition in future. Images from the various background such as forest, fields and cities can also be focused and this would help in challenging conditions like occlusion with animals partially hidden by foliage or other animal's motion blur images can be enhanced. The system focuses mostly on single object detection, however most real scenarios involve in identification of more than one species simultaneously. It is particularly common in wildlife monitoring and farm security. The model is capable of performing better on overlapping or clustered scene where more than one animal appears in the same image.

Author's Information

D.Manikandan, Assistant Professor in the Department of Computer Science and Engineering, School of Engineering, Vels Institute of Science, Technology and Advanced Studies and Part time Research Scholar in the Department of Computer Science and Engineering, College of Engineering Guindy, Anna University. He has 7 years of Research and 3 years of Teaching Experience and received Assistant Professorship by UGC NET and the National Fellowship Award from the Ministry of Social Justice and Empowerment by Government of India. His area of interest is Blockchain, Machine Learning, Deep Learning, Social Network and IoT.

M. Swarna, UG Scholar in the Department of Computer Science and Engineering- Artificial Intelligence and Machine Learning, School of Engineering, Vels Institute of Science, Technology and Advanced Studies. She has 1 year of research experience and her area of interest is Machine Learning, Deep Learning.

D. Nivetha, UG Scholar in the Department of Computer Science and Engineering- Artificial Intelligence and Machine Learning, School of Engineering, Vels Institute of Science, Technology and Advanced Studies. She has 1 year of research experience and her area of interest is Machine Learning, Deep Learning.

Dharini Kavya , UG Scholar in the Department of Computer Science and Engineering- Artificial Intelligence and Machine Learning, School of Engineering, Vels Institute of Science, Technology and Advanced Studies. She has 1 year of research experience and her area of interest is Machine Learning, Deep Learning.

Acknowledgment: The authors thanks Machine Learning laboratory, Department of Computer Science and Engineering, Vels Institute of Science, Technology and Advanced Studies (VISTAS) for providing the infrastructure support to carry out the research.

Availability of data:

All the data are provided on request

References

1. THABIT, E.M., ALSAADI, A.A., EL ABBADI, N.K.: An Automated Mammals Detection Based on SSD-MobileNet. Ibn Al-Haitham International Con ference for Pure and Applied Sciences (IHICPS), Journal of Physics: Conference Series 1879, 1–10 (2021).
2. GEETHANJALI, P., METUN, M., RAJESHWARI, M.: Advances in ecological surveillance: Real-time wildlife detection using MobileNet-SSD V2 convolutional neural network. International Journal for Research in Ap plied Science Engineering Technology (IJRASET) 11(XII), 2333–2339 (2023).
3. RAVOOR, P.C., SUDARSHAN, T.S.B., RANGARAJAN, K.: Digital Borders: De sign of an animal intrusion detection system based on deep learning. Communications in Computer and Information Science 1381, 207–218 (2021).
4. GULZAR, Y.: Fruit Image Classification Model Based on MobileNetV2 with Deep Transfer Learning Technique. Sustainability 15(3), 1906 (2023).

5. KUMAR T., GUPTA U., TANEJA V., DEEPANSHI: Animal Incursion Detection System. International Research Journal of Engineering and Technology (IRJET) 7(5), 5681–5685 (2020).
6. LI X., DU J., YANG J., LI S.: When MobileNetV2 Meets Transformer: A Balanced Sheep Face Recognition Model. Agriculture 12(8), 1126 (2022)
7. VIJAYKUMAR, A., PUJITHA, J., RACHANA, P., SHIRISHA, P., SPOORTHY, S., RAJ, N.: Animal Detection Using Deep Learning. Dogo Rangsang Research Journal, Vol-08, Issue-14, 841–845 (2021) s
8. JAYAKUMAR R, SWAMINATHAN R. HARIKUMAR, S., BANUPRIYA, N., SARANYA, S.: Animal Detection Using Deep Learning Algorithm. First International Conference on Intelligent Digital Transformation (ICIDT), Volume I, 11–15 (2019).
9. TARUN KUMAR, UTKARSH GUPTA, VANI TANEJA, DEEPANSHI: Animal Incursion Detection System. International Research Journal of Engineering and Technology (IRJET) 7(5), 5681–5686 (2020)
10. D. D. PUKALE, KOMAL DUDHANE, SHREYA DAWGHAT, ANUSHKA DESHMUKH, SHARWARI DESHMUKH: Wild Animal Detection and Identification Using Deep Learning. International Journal of Creative Research Thoughts (IJCRT) 11(7), 477–483 (2023).