Wild Animal Detection using Deep learning

Sreedevi K L
Electronics and Communication Engineering
College of Engineering Trivandrum
Thiruvananthapuram, India
klsreedevi98@gmail.com

Dr Anitha Edison
Electronics and Communication Engineering
College of Engineering Trivandrum
Thiruvananthapuram, India
anithaedison@cet.ac.in

2022 IEEE 19th India Council International Conference (INDICON) | 978-1-6654-7350-7/22/$31.00 ©2022 IEEE | DOI: 10.1109/INDICON56171.2022.10039799

***Abstract*—The collision of vehicles with animals is an emerging threat to humans and wildlife. Efficient observation of wild animals is crucial. Cost-effective techniques for observing the behaviour of wild animals are needed for both wildlife conservation and for reducing human-wildlife conflicts. Therefore an efficient system is required for wild animal detection. Since there are many different creatures, it is challenging task to manually identify each one. Animal detection can help to prevent animal-vehicle accidents and can trace animals. This will be achieved by applying effective deep learning algorithms. The objective is to create an algorithm for detecting wild animals. The depth-wise separable convolution layer, which combines point-wise and depth-wise convolution, is used to build the model. The suggested model adds zero padding in order to maintain edge characteristics and regulate the size of the output image. IWildCam data is used for testing the proposed algorithm. We achieved promising results with intersection over union value as 0.878 for detection and 99.6% accuracy for classification of wild animal.**

***Index Terms*—Wild animal, Deep learning, Depth-wise separable convolution, Zero padding, IOU**

I. INTRODUCTION

An active research area since the last many decades among wildlife researchers is the analysis of wild animals and their behavior[1]. This paper presents detection of wild animals using deep learning. This helps us to analyze wild animal habitat and behavior. This has been a challenging task[1].Since the images of wild animal appearance in cluttered background,in various light and climate conditions,some of them are in different viewpoints, and some are occluded images. Moreover, there is similarity in different class of animals[2]. Detection of wild animals is a difficult task in ecology[3].

Another reason for wild animal detection is animal-vehicle collision. It is found that each year hundreds of animals are being killed due to collisions between animals and vehicles[2]. Wild animals are extremely important for conservation and management. Animal-vehicle collision is a challenging problem as large amounts of resources got damaged and it is dangerous to human life. The number of these kinds of conflicts are increasing recently[1]. Therefore animal detection is a very important and is an emerging area.

The expansion of human population and the economic development causes over-use of natural resources, causes further changes in ecosystems[4]. This affects wildlife population, habitat and behavior[1]. Detection of wild animals, therefore, is essential for conservation and management decisions for a balanced and sustainable ecosystems in accordance with those changes[5].

Numerous amount of data on wildlife are available nowadays. Camera-trap procedure and different technologies are used in analysing wildlife which are cheap and ready to use[1]. The analysis related to wildlife has become more convenient with the expansion in data on wildlife. IWildCam dataset is a new camera trap dataset[6]. This dataset contain images collected from 143 locations . The classes include tiger, raccoon,cheetah, wild dog, rodent, deer, fox, lion , etc.[6].

The knowledge from camera images shows ththa image processing is needed. By knowledge analysis, various techniques are available for wild animal detection such as using linear support vector machine [7], convolutional neural network (CNN) model [7], or fine-tuning CNN models model weights pretrained on a very large dataset like ImageNet [8]. These approaches considers of wildlife detection as a major factor. However, this lead to various challenges such as the accuracy of an automated wildlife monitoring in practice needs a consideration[9]. For image detection , an enormous amount of pre-processing is required on input images for detecting and finding animals in an image [10]. The inaccurate performance obtained by the system monitoring, eventhough the use of complete automation, shows that much more improvements are needed for practical application [11].

It is crucial for the efficient analysis of wild animals in their natural surroundings . Since there are huge number of animals, manually analyzing them is a difficult task. Animal vehicle collision is an emerging threat to humans and wildlife. Conservation of wildlife and the considering the conflicts needs cost less process for monitoring wild animals[12]. Therefore a efficient system is required for wild animal detection. Animal detection help to prevent animal-vehicle accidents and can trace animal[13]. By applying effective deep learning algorithms this can be achieved. Therefore we develop

978-1-6654-7350-7/22/$31.00 ©2022 IEEE

an algorithm for wild animal detection.

II. Related Works

Detection of animal is an important research area since several decades. Many research is going on to develop a system for conservation of wildlife and monitoring wildlife human conflict.

Wilber et al.[14] proposed a lightweight procedure for detection of wild animal by using Support-vector machine (SVM). In their work images captured periodically from the video . Animals within the video can be detected and classification is done. They also analyzed that most of the previous works requires a various computing resources. Also they use inappropriate conditions that limit the process and cause various issues. They found out that these are not suited for fauna . They come to the solution that it could not achieve perfection in accuracy eventhough its species can be correctly identified .

Badrinath et al[15]. proposed a system for a surveillance of realife. The purpose is to distinguish dangerous animals. For feature extraction, they used SVM and histogram of oriented gradients (HOG) . Unfortunately, the tracking component details of the system was not given, they limit to notify on its importance as existing surveillance only when an animal enters the scene.

Khan et al[16] proposed that the various methods of machine learning can be used for detection. The enormous amount of data makes deep learning more powerful method. The computer vision include image classification and object detection. Image classification categorizes the input images. When multiple images are present, object detection can be used to classify and detect images.

III. METHODOLOGY

A. Architecture

The input image is made uniform to (125,125,3). The topmost layer is standard convolution layer. The number of filters of standard convolution layer is 64 and kernel size is 3*3. The standard convolution layer is then there is Batch Normalization and ReLu. The activation used is ReLu. Convolution layers extracts features from the image using different numbers of kernels[3]. Batch normalization normalizes the output of the previous layers. It can avoid over fitting of the model. Rectified Linear Unit(ReLu) helps to stop the exponential growth in the computation operations in the neural network. The depth-wise separable convolution layer is the second layer. Both the point wise convolution and the depth wise convolution are included. Depth wise convolution has kernel size (3,3), strides=1. Point wise convolution has kernel size (1,1), strides=1. Series of such layers are added. Following each of them there are Batch Normalization and ReLU. In this proposed model, zero padding is also used. It is used to preserve the original input size[16]. In order to maintain edge characteristics and adjust the size of the output image, zero padding is used[17]. The typical pooling kernel size is (4,4). In order to generate a down sampled feature map, it computes the average value for patches of a feature map[18]. A fully connected layer is the following layer in which all activation unit of the next layer is connected to the inputs from the previous layer[20]. Softmax activation is employed. Fig.1 denote architecture of the created model.

After creating the model , it is trained with training dataset. For training the model epoch value is given as 10 and the dropout value is given as 0.2. The width of the network, alpha is given as 0.5. Detection of wild animal in an image is shown in bounding boxes.

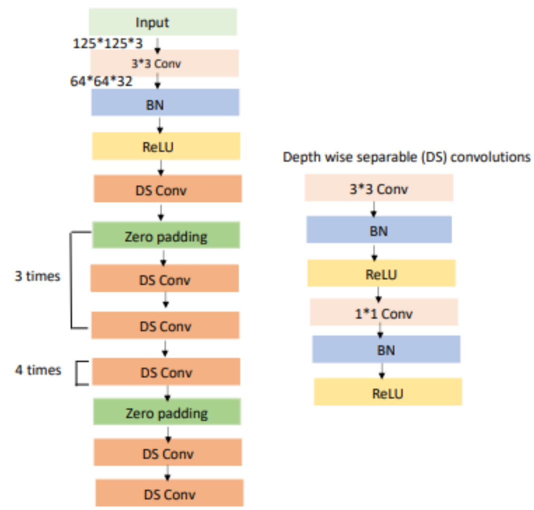

Fig. 1. Architecture of the proposed model.

B. Data collection

Dataset used is IWildCam. The datasets is collected from Kaggle. IWildCam Dataset is a new camera trap dataset which contain images collected from 143 locations[6]. The classes include tiger, raccoon, wild dog, deer, fox, mountain lion, cheetah. Some of the sample images from the IWildCam dataset is shown in Fig.2.

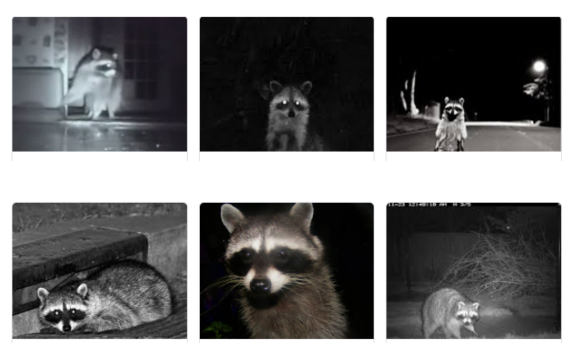

Fig. 2. Sample images from IWildCam dataset.

C. Evaluation Metrics

1) Intersection over union(IOU):

Intersection over Union (IOU) is an analysis parameter for object detection. It is the most popular evaluation metric. Its value ranges from 0 to 1 [21]. The ground truth bounding box and the predicted bounding box overlap is specified . The absence of any overlap between the boxes is indicated by an IOU value of 0. The bounding boxes fully overlap when the IOU has a value of 1[22]. The expression for intersection over union is shown below :

$$IOU = \frac{Areaofintersection}{Areaofunion} \tag{1}$$

To determine whether a detection is a true or a false positive, a threshold value is given. It is considered as true positive if the IOU value exceeds the threshold value otherwise, it is considered as false positive[22]. In this work threshold value is set as 0.5. An Intersection over union value larger than 0.5 is considered a good prediction.

2) *Precision and Recall:*

Precision quantifies how precise the predictions are and is measured in percentage . It measures how many of the positive predictions made are correct[23]. The expression for precision is shown as :

$$Precision = \frac{Truepositive}{Truepositive + Falsepositive} \tag{2}$$

Recall measures however well the model detects all the positives. It is a measure of how many of the positive cases the model correctly predicted, over all the positive cases within the data. The expression for recall is shown below :

$$Recall = \frac{Truepositive}{Truepositive + Falsenegative} \tag{3}$$

3) *F1 score:*

F1 score is determined by the mean of precision and recall. When the F1 score is 1, the model is considered to be perfect , when it is 0, the model is not considered to be perfect. Below is a representation of the F1 score :

$$IOU = 2\frac{Precision * Recall}{Precision + Recall} \tag{4}$$

4) *Confusion matrix:*

A table that depicts the general performance of a classification method is called a confusion matrix. The effectiveness of a classification system is shown and summarised using a confusion matrix[24].

IV. Experimental Results

Necessary libraries such as numpy, pandas, keras, matplotlib are imported. Sample detection result obtained as shown in Fig.3 and Fig.4. The results shows that the proposed system is able to detect wild animals correctly. All images were resized to 125x125 pixels. Bounding boxes are drawn around the animals to detect the presence of wild animal[25].

Alpha, width multiplier that is used to controls the width of the network. If alpha is much less than 1.0, it proportionally decreases the number of filters in every layer. If

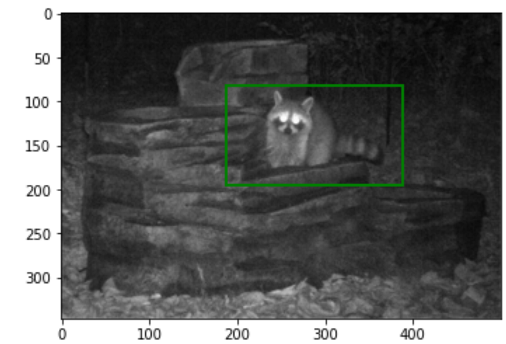

Fig. 3. Sample detection results of IWildCam dataset.

Fig. 4. Sample detection results of IWildCam dataset.

alpha more than 1.0, it will increase the number of filters in every layer. The intersection over union is calculated for different values of alpha such as 1, 0.75, 0.5 and 0.25 and the IOU value obtained as 0.855, 0.84, 0.878 and 0.853 respectively for each values of alpha. It is identified that with alpha value 0.5 highest IOU value is obtained. IOU value is calculated for various values of alpha and is obtained as shown in the TABLE 1. With dropout value 0.5 and epoch 10, the accuracy and loss rate is obtained. Accuracy and loss rate curve is plotted as shown in the Fig.5 and Fig.6 respectively. It is seen that training and validation accuracy will increase with epoch. Whereas training and validation loss decreases with epoch. The accuracy of the proposed model is 99.6%. The test Loss obtained as 0.0066 and test Accuracy is obtained as 0.996. This shows the proposed model is effective in detection and classification of wild animals.

In order to calculate precision, recall and F1 score three class of wild animal is considered. Class 0 belong to tiger,

TABLE I
IOU VALUE OBTAINED FOR DIFFERENT VALUES OF ALPHA

ALPHA	IOU
1	0.855
0.75	0.84
0.5	0.878
0.25	0.853

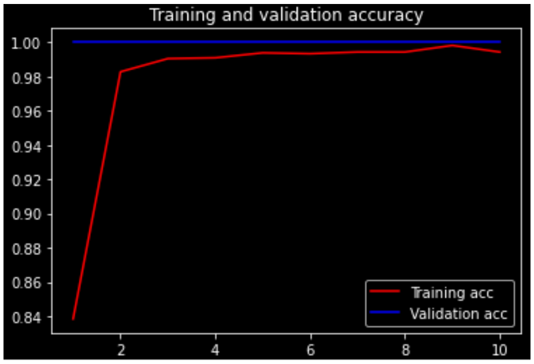

Fig. 5. Training and Validation Accuracy.

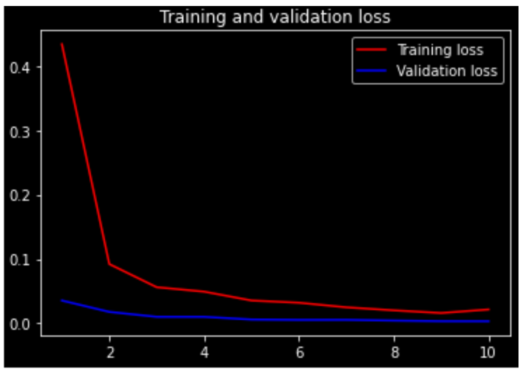

Fig. 6. Training and Validation Loss.

class 1 belongs to raccoon and class 2 belongs to cheetah. Each of 100 images of theses three classes are taken. Precision, recall and F1 score of tiger is obtained as 1.00, 0.99, 0.99 respectively. Precision, recall and F1 score of raccoon is obtained as 1.00, 1.00, 1.00 respectively. Precision, recall and F1 score of cheetah is obtained as 0.99, 1.00, 1.00 respectively. The average of F1 score is 1.00. Fig.7 shows Precision, Recall and F1 score values obtained. These results shows that our model is efficient.

	Precision	Recall	F1-score	Number of images
Tiger	1.00	0.99	0.99	100
Raccoon	1.00	1.00	1.00	100
Cheetah	0.99	1.00	1.00	100
Accuracy			1.00	300
Macro average	1.00	1.00	1.00	300
Weighted average	1.00	1.00	1.00	300

Fig. 7. Precision, Recall and F1 score.

The Confusion matrix is obtained by considering three classes of wild animals such as tiger, raccoon and cheetah as shown in the Fig.8. The three class of wild animals tiger, raccoon, cheetah is represented as 0, 1 and 2. By the analysis of confusion matrix, it is understood that out of 100 images of tiger represented as class 0, 99 images are correctly classified and one image is wrongly classified as cheetah. All the 100 images of Raccoon, which is represented as class 1 is correctly classified. All the 100 images of cheetah ,which is represented as class 2 is also correctly classified. These end results indicates that our proposed approach is powerful in classifying wild animals.

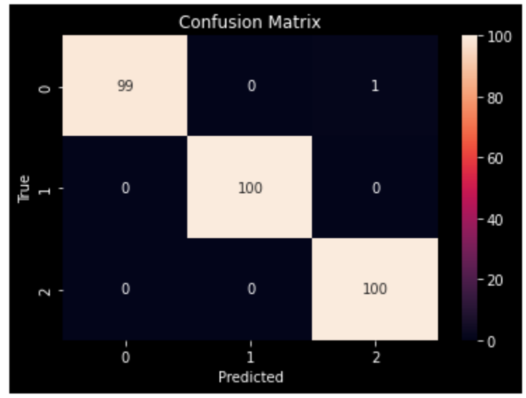

Fig. 8. Confusion matrix for classification.

Conclusion

In this paper, a method for detection of wild animals is proposed. Our algorithm is effective in detection and to identify the class of wild animals. The model is developed by using depth wise separable convolution layer. Zero padding is added in the proposed model to preserve features that exist at edges and to control the output image size. This is performed over the IWildCam data set which is an open source data set. The end results indicates that that the proposed approach is powerful in detection and classification of wild animals. The intersection over union value obtained is 0.86 which is greater than 0.5 which is a good prediction. With this system the predicted bounding box is plotted clearly. The accuracy and loss rate curve is plotted. Finally the evaluation parameter for classification is measured and shows that the result is accurate. The performance can be improved by multiple sensor, LADAR or IR sensors.

References

[1] S. Matuska, R. Hudec, M. Benco, P. Kamencay and M.Zachariasova, “A novel system for automatic detection and classification of animal,” 2014 ELEKTRO 2018, pp. 76-80.

[2] Sanjay Santhanam, Sudhir Sidhaarthan B, Sai Sudha Panigrahi, “Animal Detection for Road safety using Deep Learning”, 2021 International conference on computational intelligence and computing applications, 2021.

[3] Z. Zhang, Z. He, G. Cao and W. Cao,”Animal Detection From Highly Cluttered Natural Scenes Using Spatiotemporal Object Region Proposals and Patch Verification” IEEE Transactions on Multimedia ,vol. 18, no. 10, pp. 2079-2092, Oct. 2016.

[4] N. Li, W. Kusakunniran and S. Hotta, ”Detection of Animal Behind Cages Using Convolutional Neural Network,” 2020 17th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON), 2020, pp. 242-245, doi: 10.1109/ECTI-CON49241.2020.9158137.

[5] A. V. P. Ram, A. Sarath Prakash, A. I. Ahamed, K. Anirudh, M. M. Arvindh and N. Nithyavathy, ”A Self Induced Warning System for Wild Animal Trespassing Using Machine Vision System,” 2018 International Conference on Intelligent Computing and Communication for Smart World (I2C2SW), 2018, pp. 349-353, doi: 10.1109/I2C2SW45816.2018.8997520.

[6] Sara Beery , Dan Morris, Pietro Perona, “ The iWildCam 2019 Challenge Dataset”, 2019 .

[7] Y. Munian, A. Martinez-Molina and M. Alamaniotis, ”Intelligent System for Detection of Wild Animals Using HOG and CNN in Automobile Applications,” 11th International Conference on Information, Intelligence, Systems and Applications, 2020.

[8] Nagaraju Andavarapu and Valli Kumari Vatsavayi,”Wild-animal recognition in agriculture farms using W-COHOG for agro-security”, International Journal of Computational Intelligence Research, doi: 10.1109/ECTI-CON49241.2020.9158137, vol. 13, no. 9, pp. 2247-2257, 2017.

[9] R. Shanmugasundaram and S. Pavithra, ”IOT based animal tracking and monitoring system in Zoo” South Asian Journal of Engineering and Technology ,vol. 3, no. 2, 2017.

[10] R. Sato, H. Saito, Y. Tomioka and Y. Kohira, ”Energy Reduction Methods for Wild Animal Detection Devices,” in IEEE Access, vol. 10, pp. 24149-24161, 2022, doi: 10.1109/ACCESS.2022.3155242.

[11] Mohammad Sadegh Norouzzadeh, Anh Nguyen, Margaret Kosmala, ”Automatically identifying counting and describing wild animals in camera-trap images with deep learning” Proc. of the National Academy of Sciences, vol. 115, no. 25, pp. E5716-E5725, 2018

[12] Y. A. Roopashree, M. Bhoomika, R. Priyanka, K. Nisarga and S. Behera, ”Monitoring the Movements of Wild Animals and Alert System using Deep Learning Algorithm,” 2021 International Conference on Recent Trends on Electronics, Information, Communication Technology (RTEICT), 2021, pp. 626-630.

[13] S. Santhanam, S. S. B, S. S. Panigrahi, S. K. Kashyap and B. K. Duriseti, ”Animal Detection for Road safety using Deep Learning,” 2021 International Conference on Computational Intelligence and Computing Applications (ICCICA), 2021, pp. 1-5, doi: 10.1109/ICCICA52458.2021.9697287.

[14] M. Wilber et al. ”Animal recognition in the mojave desert: Vision tools for field biologists”, IEEE Workshop on Applications of Computer Vision (WACV), pp. 206-213, 2013.

[15] N.M. Badrinath, N. Bharat, S. Madhumathi, and B.S. Kariyappa. ”Pattern recognition using video surveillance for wildlife applications”, International Journal of Electronics and Communication Engineering Research and Development (IJECERD),Volume 4, 2014.

[16] Mahdi Hashemi, “Enlarging smaller images before inputting into convolutional neural network: zero-padding vs. interpolation”, Journal of Big Data , doi.org/10.1186/s40537-019-0263-7, 2019.

[17] H. Li, M. Lu and Z. Feng, ”Improved Zero-Padding Method for Rapid Long PN-Code Acquisition,” in IEEE Transactions on Signal Processing, vol. 56, no. 8, pp. 3795-3799, Aug. 2008, doi: 10.1109/TSP.2008.924859.

[18] A. Khan, A. Sohail, U. Zahoora, and A. S. Qureshi ”A survey of the recent architectures of deep convolutional neural networks”, Artificial Intelligence Review, Apr 2020.

[19] T. Saito and M. Akiyama, ”A Study of Image Recognition Rate with and without Padding of 2D Convolutional Layer for Deep Learning in CanSat,” 2021 IEEE 10th Global Conference on Consumer Electronics (GCCE), 2021, pp. 816-817, doi: 10.1109/GCCE53005.2021.9621978.

[20] S. Schneider, G. W. Taylor, and S. C. Kremer ”Deep learning object detection methods for ecological camera trap data”, pp. 321-328. doi: 10.1109/CRV.2018.00052, 2018.

[21] H. Rezatofighi, N. Tsoi, J. Gwak, A. Sadeghian, I. Reid and S. Savarese, ”Generalized Intersection Over Union: A Metric and a Loss for Bounding Box Regression,” 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2019, pp. 658-666, doi: 10.1109/CVPR.2019.00075.

[22] Z. Qin, Q. Li, H. Li, X. Dong and Z. Ren, ”Advanced Intersection over Union Loss for Visual Tracking,” 2019 Chinese Automation Congress (CAC), 2019, pp. 5869-5873, doi: 10.1109/CAC48633.2019.8997311.

[23] Y. Zhang et al., ”Detecting third-party libraries in Android applications with high precision and recall,” 2018 IEEE 25th International Conference on Software Analysis, Evolution and Reengineering (SANER), 2018, pp. 141-152, doi: 10.1109/SANER.2018.8330204.

[24] F. J. Ariza-Lopez, J. Rodriguez-Avi and M. V. Alba-Fernandez, ”Complete Control of an Observed Confusion Matrix,” IGARSS 2018 - 2018 IEEE International Geoscience and Remote Sensing Symposium, 2018, pp. 1222-1225, doi: 10.1109/IGARSS.2018.8517540.

[25] J. Parham, C. Stewart, J. Crall, D. Rubenstein, J. Holmberg and T. Berger-Wolf, ”An Animal Detection Pipeline for Identification,” 2018 IEEE Winter Conference on Applications of Computer Vision (WACV), 2018, pp. 1075-1083.