ISSN: 2454-9940

INTERNATIONAL JOURNAL OF APPLIED SCIENCE ENGINEERING AND MANAGEMENT

E-Mail :
editor.ijasem@gmail.com
editor@ijasem.org

www.ijasem.org

INTERNATIONAL JOURNAL OF APPLIED SCIENCE ENGINEERING AND MANAGEMENT
https://doi.org/10.5281/zenodo.14350855
ISSN 2454-9940
www.ijasem.org
Vol 18, Issue 4, 2024

Creating Alert Messages Based On Wild Animal Activity Detection Using Hybrid Deep Neural Networks

Nomula Sowmya[1], Dr.B.Srinivas Rao[2]

[1]PG Scholar, Department of CSE, Teegala Krishna Reddy Engineering College (Autonomous Institution), Medbowli, Meerpet, Saroornagar, Hyderabad
[2] Professor, Department of CSE, Teegala Krishna Reddy Engineering College (Autonomous Institution), Medbowli, Meerpet, Saroornagar, Hyderabad

ABSTRACT

The issue of animal attacks is increasingly concerning for rural populations and forestry workers. To track the movement of wild animals, surveillance cameras and drones are often employed. However, an efficient model is required to detect the animal type, monitor its locomotion and provide its location information. Alert messages can then be sent to ensure the safety of people and foresters. While computer vision and machine learning-based approaches are frequently used for animal detection, they are often expensive and complex, making it difficult to achieve satisfactory results. This paper presents a Hybrid Visual Geometry Group (VGG)−19+ Bidirectional Long Short-Term Memory (Bi-LSTM) network to detect animals and generate alerts based on their activity. These alerts are sent to the local forest office as a Short Message Service (SMS) to allow for immediate response. The proposed model exhibits great improvements in model performance, with an average classification accuracy of 98%, a mean Average Precision (mAP) of 77.2%, and a Frame Per Second (FPS) of 170. The model was tested both qualitatively and quantitatively using 40,000 images form three different bench mark datasets with 25 classes and achieved a mean accuracy and precision of above 98%. This model is a reliable solution for providing accurate animal-based information and protecting human lives.

Keywords:Visual Geometry Group (VGG),Bidirectional Long Short-Term Memory (Bi-LSTM),Short Message Service (SMS)

I. INTRODUCTION

In general, animal activity detection creates numerous challenges for researchers due to the continuous streaming of inputs and the cluttered backgrounds. There are huge varieties of wildlife categories with different facial, nose, body, and tail structures. The detection and classification of such animals in video sequences and the processing of huge feature maps demand the need to develop a robust framework. Such developments in real-time cases need large-scale video data for training and testing purposes and

INTERNATIONAL JOURNAL OF APPLIED SCIENCE ENGINEERING AND MANAGEMENT
https://doi.org/10.5281/zenodo.14350855
ISSN 2454-9940
www.ijasem.org
Vol 18, Issue 4, 2024

high GGPU-based computing resources. Moreover, the incorporating techniques should handle the data in an intelligent way to produce plausible results. Hence, there is a high demand for developing such a model to detect animal activities in forest regions. Although numerous advancements have been made in this technological era, research in this area still seeks higher attention to produce a strong model. With this work, we can save humans from sudden animal attacks as well as send alert messages with location information to the forest officers for quick action.

The goal of this project is to develop a system that can instantly identify animal detection and classify what type of animal it is. We will employ neural network methods. Via a through CNN techniques and using Hybrid VGG-19+Bi-LSTM models, the study will show highest accuracy animal detection

The main objectives of these systems are offering better monitoring services and help to find the activities of animals and detect if there is any hunting by humans or hindrance to wildlife. These clusters of activities, such as tracking the animal object and finding its activity and generating the alert messages, pose huge complexity in the Deep Learning area. Recent developments in Deep Learning techniques have produced impressive results in image recognition, classification, and generation tasks. Due to these developments, we focus our aim on developing a robust model for monitoring the activities of animals and generating alerts to the forest officers in case of any abnormal activity such as hunting, animals entering into human living areas or agricultural land. The development of the proposed model investigates this problem from multiple angles to provide a better solution.

II. RELATED WORK

Animal detection using deep learning algorithm

Authors: N. Banupriya, S. Saranya, R. Swaminathan, S. Harikumar, and S. Palanisamy

Abstract: Efficient and reliable monitoring of wild animals in their natural habitat is essential. This project develops an algorithm to detect the animals in wild life. Since there are large number of different animals manually identifying them can be a difficult task. This algorithm classifies animals based on their images so we can monitor them more efficiently. Animal detection and classification can help to prevent animal-vehicle accidents, trace animals and prevent theft. This can be achieved by applying effective deep learning algorithms. Experimental results indicate that our proposed technique is significantly better than other techniques used for wild animal activity detection.

https://doi.org/10.5281/zenodo.14350855
ISSN 2454-9940
www.ijasem.org
Vol 18, Issue 4, 2024

Z. Zhang, Z. He, G. Cao, and W. Cao, ''Animal detection from highly cluttered natural scenes using spatiotemporal object region proposals and patch verification, 2016

Authors: Z. Zhang, Z. He, G. Cao, and W. Cao,

Abstract: In this paper, we consider the animal object detection and segmentation from wildlife monitoring videos captured by motion-triggered cameras, called camera-traps. For these types of videos, existing approaches often suffer from low detection rates due to low contrast between the foreground animals and the cluttered background, as well as high false positive rates due to the dynamic background. To address this issue, we first develop a new approach to generate animal object region proposals using multilevel graph cut in the spatiotemporal domain. We then develop a cross-frame temporal patch verification method to determine if these region proposals are true animals or background patches. We construct an efficient feature description for animal detection using joint deep learning and histogram of oriented gradient features encoded with Fisher vectors. Our extensive experimental results and performance comparisons over a diverse set of challenging camera-trap data demonstrate that the proposed spatiotemporal object proposal and patch verification framework outperforms the state-of-the-art methods, including the recent Faster-RCNN method, on animal object detection accuracy by up to only about 4.5%.

Animal Recognition and Identification. with Deep Convolutional Neural Networks for Automated. Wildlife Monitoring.

Authors: H. Nguyen et al.

Abstract: Efficient and reliable monitoring of wild animals in their natural habitats is essential to inform conservation and management decisions. Automatic covert cameras or "camera traps" are being an increasingly popular tool for wildlife monitoring due to their effectiveness and reliability in collecting data of wildlife unobtrusively, continuously and in large volume. However, processing such a large volume of images and videos captured from camera traps manually is extremely expensive, time-consuming and also monotonous. This presents a major obstacle to scientists and ecologists to monitor wildlife in an open environment. Leveraging on recent advances in deep learning techniques in computer vision, we propose in this paper a framework to build automated animal recognition in the wild, aiming at an automated wildlife monitoring system. In particular, we use a single-labelled dataset from Wildlife Spotter project, done by citizen scientists, and the state-of-the-art deep convolutional neural network architectures, to train a computational system capable of filtering

INTERNATIONAL JOURNAL OF APPLIED SCIENCE ENGINEERING AND MANAGEMENT
https://doi.org/10.5281/zenodo.14350855
ISSN 2454-9940
www.ijasem.org
Vol 18, Issue 4, 2024

animal images and identifying species automatically. Our experimental results achieved an accuracy at 96.6% for the task of detecting images containing animal, and 90.4% for identifying the three most common species among the set of images of wild animals taken in South-central Victoria, Australia, demonstrating the feasibility of building fully automated wildlife observation. This, in turn, can therefore speed up research findings, construct more efficient citizen science-based monitoring systems and subsequent management decisions, having the potential to make significant impacts to the world of ecology and trap camera images analysis. However, this experiment results are time consuming to detect the wild animal activity detection.

Wild Animal Detection Using Deep Convolutional Neural Network.

Authors: Verma, Gyanendra & Gupta, Pragya.

Abstract: Wildlife monitoring and analysis are an active research field since last many decades. In this paper, we focus on wildlife monitoring and analysis through animal detection from natural scenes acquired by camera-trap networks. The image sequences obtained from camera-trap consist of highly cluttered images that hinder the detection of animal resulting in low-detection rates and high false discovery rates. To handle this problem, we have used a camera-trap database that has candidate animal proposals using multilevel graph cut in the spatiotemporal domain. These proposals are used to create a verification phase that identifies whether a given patch is animal or background. We have designed animal detection model using self-learned Deep Convolutional Neural Network (DCNN) features. This efficient feature set is then used for classification using state-of-the-art machine learning algorithms, namely support vector machine, k-nearest neighbour, and ensemble tree. Our intensive results show that our detection model using DCNN features provides higher accuracy of on standard camera-trap dataset. Although, this model provides high accuracy it cannot work properly under cluttered backgrounds

Wild Animal Intrusion Detection System using YOLO.

Authors: Aibin Abraham, Bibin Mathew.

Abstract: Agriculture plays a crucial role in the economy, and farmers strive to increase their crop yields annually. Hence effective reconnaissance is vital for farmlands and rural terrains to prevent unauthorized access and protect crops from animal damage. The expansion of agricultural lands into wildlife territories has escalated human-wildlife conflicts, with crop destruction by animals becoming a major concern. To address this, our project proposes an alerting system using

INTERNATIONAL JOURNAL OF APPLIED SCIENCE ENGINEERING AND MANAGEMENT
https://doi.org/10.5281/zenodo.14350855
ISSN 2454-9940
www.ijasem.org
Vol 18, Issue 4, 2024

YOLOv3, a real- time object detection algorithm based on deep convolutional neural networks, to classify and monitor animals that intrude into agricultural areas. This algorithm enables efficient identification and tracking of animals, aiding in mitigating crop damage and ensuring the preservation of wildlife in their natural habitats. Whenever an animal is detected, the system will send an SMS to the landowner and forest officials, providing them with early warning notifications to take appropriate actions based on the intruder's type. This proposed system offers significant benefits to farmers, helping them increase yields and protect both humans and livestock from wild animal attacks.

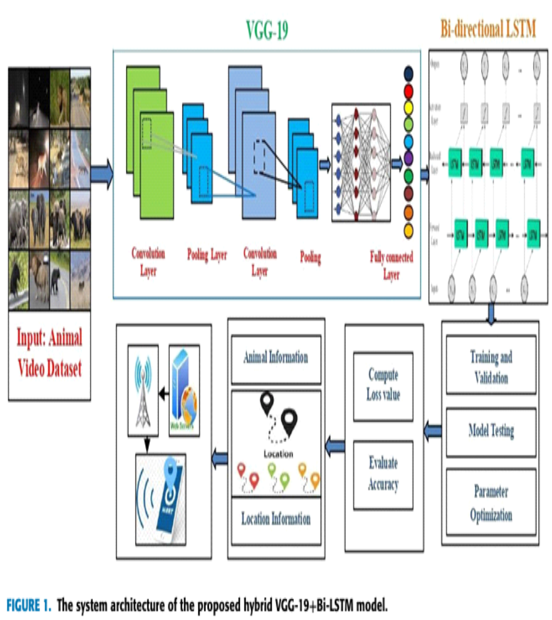

Fig.1 System Architecture.

III.IMPLEMENTATION

1. Service Provider:

In this module, the Service Provider has to login by using valid user name and password. After login successful he can do some operations such as Browse and Train & Test Data Sets, View Trained and Tested Accuracy in Bar Chart, View Trained and Tested Accuracy Results, View Prediction Of Animal Activity Detection Type, View Animal Activity Detection Type Ratio, Download Predicted Data Sets, View Animal Activity Detection Type Ratio Results, View All Remote Users.

2. View and Authorize Users

In this module, the admin can view the list of users who all registered. In this, the admin can view the user's details such as, user name, email, address and admin authorize the users.

3. Remote User

In this module, there are n numbers of users are present. User should register before doing any operations. Once user registers, their details will be stored to the database. After registration successful, he has to login by using authorized user name and password. Once Login is successful user will do some operations like REGISTER AND LOGIN, PREDICT ANIMAL ACTIVITY DETECTION TYPE, VIEW YOUR PROFILE.

INTERNATIONAL JOURNAL OF APPLIED SCIENCE ENGINEERING AND MANAGEMENT

https://doi.org/10.5281/zenodo.14350855

ISSN 2454-9940

www.ijasem.org

Vol 18, Issue 4, 2024

IV.ALGORITHMS

Decision tree classifiers:

Decision tree classifiers are used successfully in many diverse areas. Their most important feature is the capability of capturing descriptive decision-making knowledge from the supplied data. Decision tree can be generated from training sets. The procedure for such generation based on the set of objects (S), each belonging to one of the classes C1, C2, …, Ck is as follows:

Step 1. If all the objects in S belong to the same class, for example Ci, the decision tree for S consists of a leaf labelled with this class

Step 2. Otherwise, let T be some test with possible outcomes O1, O2,…On. Each object in S has one outcome for T so the test partitions S into subsets S1, S2,… Sn where each object in Si has outcome Oi for T. T becomes the root of the decision tree and for each outcome Oi we build a subsidiary decision tree by invoking the same procedure recursively on the set Si.

Gradient boosting:

Gradient boosting is a machine learning technique used in regression and classification tasks, among others. It gives a prediction model in the form of an ensemble of weak prediction models, which are typically decision trees. When a decision tree is the weak learner, the resulting algorithm is called gradient-boosted trees; it usually outperforms random forest. The gradient-boosted trees model is built in a stage-wise fashion as in other boosting methods, but it generalizes the other methods by allowing optimization of an arbitrary differentiable loss function.

K-Nearest Neighbours (KNN):

- Simple, but a very powerful classification algorithm
- Classifies based on a similarity measure
- Non-parametric
- Lazy learning
- Does not "learn" until the test example is given

Whenever we have a new data to classify, we find its K-nearest neighbours from the training data. Example

- Training dataset consists of k-closest examples in feature space
- Feature space means, space with categorization variables (non-metric variables)

Learning based on instances, and thus also works lazily because instance close to the input vector for test or prediction may take time to occur in the training dataset

Logistic regression Classifiers:

Logistic regression analysis studies the association between a categorical dependent variable and a set of independent (explanatory) variables. The name logistic regression is used when the dependent variable has only two values, such as 0 and 1 or

INTERNATIONAL JOURNAL OF APPLIED SCIENCE ENGINEERING AND MANAGEMENT
https://doi.org/10.5281/zenodo.14350855
ISSN 2454-9940
www.ijasem.org
Vol 18, Issue 4, 2024

Yes and No. The name multinomial logistic regression is usually reserved for the case when the dependent variable has three or more unique values, such as Married, Single, Divorced, or Widowed. Although the type of data used for the dependent variable is different from that of multiple regression, the practical use of the procedure is similar.

This program computes binary logistic regression and multinomial logistic regression on both numeric and categorical independent variables. It reports on the regression equation as well as the goodness of fit, odds ratios, confidence limits, likelihood, and deviance. It performs a comprehensive residual analysis including diagnostic residual reports and plots. It can perform an independent variable subset selection search, looking for the best regression model with the fewest independent variables. It provides confidence intervals on predicted values and provides ROC curves to help determine the best cutoff point for classification. It allows you to validate your results by automatically classifying rows that are not used during the analysis.

Naïve Bayes:

The naive bayes approach is a supervised learning method which is based on a simplistic hypothesis: it assumes that the presence of a particular feature of a class is unrelated to the presence of any other feature. Yet, despite this, it appears robust and efficient. Its performance is comparable to other supervised learning techniques. Various reasons have been advanced in the literature. In this tutorial, we highlight an explanation based on the representation bias. The naive bayes classifier is a linear classifier, as well as linear discriminant analysis, logistic regression or linear SVM. The difference lies on the method of estimating parameters of the classifier.While the Naive Bayes classifier is widely used in the research world, it is not widespread among practitioners which want to obtain usable results. On the one hand, the researchers found especially it is very easy to program and implement it, its parameters are easy to estimate, learning is very fast even on very large databases, its accuracy is reasonably good in comparison to the other approaches. On the other hand, the final users do not obtain a model easy to interpret and deploy, they do not understand the interest of such a technique.Thus, we introduce in a new presentation of the results of the learning process. The classifier is easier to understand, and its deployment is also made easier. In the first part of this tutorial, we present some theoretical aspects of the naive bayes classifier. Then, we implement the approach on a dataset with Tanagra. We compare the obtained results (the parameters of the model) to those obtained with other linear approaches such as the logistic regression, the linear discriminant analysis and the linear SVM.

INTERNATIONAL JOURNAL OF APPLIED SCIENCE ENGINEERING AND MANAGEMENT
https://doi.org/10.5281/zenodo.14350855
ISSN 2454-9940
www.ijasem.org
Vol 18, Issue 4, 2024

We note that the results are highly consistent. This largely explains the good performance of the method in comparison to others. In the second part, we use various tools on the same dataset.

Random Forest:

Random forests or random decision forests are an ensemble learning method for classification, regression and other tasks that operates by constructing a multitude of decision trees at training time. For classification tasks, the output of the random forest is the class selected by most trees. For regression tasks, the mean or average prediction of the individual trees is returned. Random decision forests correct for decision trees' habit of overfitting to their training set. Random forests generally outperform decision trees, but their accuracy is lower than gradient boosted trees. However, data characteristics can affect their performance. The first algorithm for random decision forests was created in 1995 by Tin Kam Ho using the random subspace method, which, in Ho's formulation, is a way to implement the & quot;stochastic discrimination " approach to classification proposed by Eugene Kleinberg. An extension of the algorithm was developed by Leo Breiman and Adele Cutler, who registered "Random Forests" as a trademark in 2006 (as of 2019, owned by Minitab, Inc.).The extension combines Breiman's "bagging" idea and random selection of features, introduced first by Ho[1] and later independently by Amit and Geman[13] in order to construct a collection of decision trees with controlled variance. Random forests are frequently used as "blackbox" models in businesses, as they generate reasonable predictions across a wide range of data while requiring little configuration.

SVM:

In classification tasks a discriminant machine learning technique aims at finding, based on an independent and identically distributed training dataset, a discriminant function that can correctly predict labels for newly acquired instances. Unlike generative machine learning approaches, which require computations of conditional probability distributions, a discriminant classification function takes a data point x and assigns it to one of the different classes that are a part of the classification task. Less powerful than generative approaches, which are mostly used when prediction involves outlier detection, discriminant approaches require fewer computational resources and less training data, especially for a multidimensional feature space and when only posterior probabilities are needed. From a geometric perspective, learning a classifier is equivalent to finding the equation for a multidimensional surface that best separates the different classes in the feature space.

Hybrid Visual Geometry Group (VGG)−19+:

INTERNATIONAL JOURNAL OF APPLIED SCIENCE ENGINEERING AND MANAGEMENT
https://doi.org/10.5281/zenodo.14350855
ISSN 2454-9940
www.ijasem.org
Vol 18, Issue 4, 2024

VGG-19 is a deep convolutional neural network with 19 weight layers, comprising 16 convolutional layers and 3 fully connected layers. The architecture follows a straightforward and repetitive pattern, making it easier to understand and implement.

VGG-19: This is a convolutional neural network architecture that was proposed as part of the Visual Geometry Group's work. It consists of 19 layers (hence the name VGG-19) and was designed for image classification tasks.

Hybrid: The term "hybrid" suggests that modifications or enhancements have been made to the original VGG-19 architecture. These modifications could include changes in the layers, introduction of new components, or adaptations for specific tasks.

Usage: Researchers often create "hybrid" models by integrating ideas or techniques from different sources to improve performance or adapt the model for specialized tasks. Therefore, "Hybrid VGG-19+" could imply a version of VGG-19 that has been customized or enhanced in some way to better suit a particular application or to achieve better performance metrics.

Bidirectional Long Short-Term Memory (Bi-LSTM) network:

A Bidirectional Long Short-Term Memory (Bi-LSTM) network is a type of recurrent neural network (RNN) architecture that enhances the capability of standard LSTMs by processing input data in both forward and backward directions. In a Bi-LSTM, the input sequence is processed in two ways: one in forward order and another in backward order. The bidirectional nature allows the model to have a more comprehensive view of the sequence, incorporating both past and future contexts into the current prediction or classification decision.

V.RESULTS

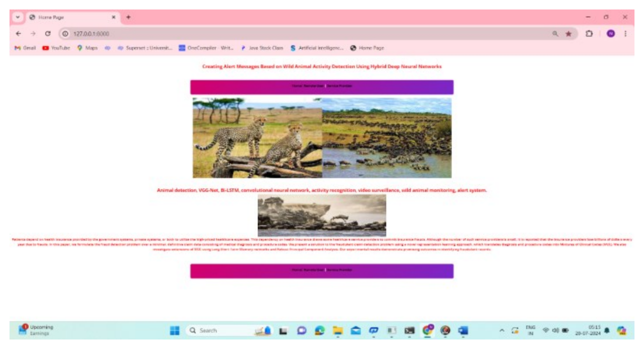

Fig 1: Home Page.

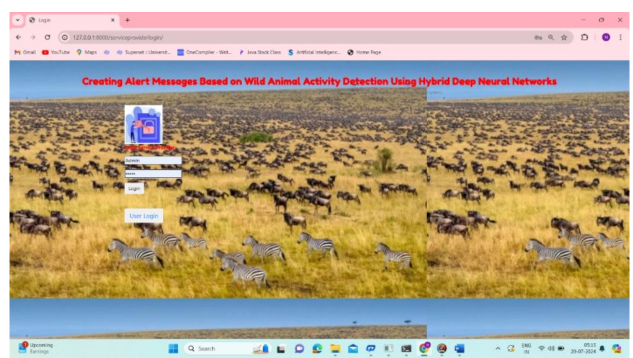

Fig 2: Admin Login.

INTERNATIONAL JOURNAL OF APPLIED SCIENCE ENGINEERING AND MANAGEMENT
https://doi.org/10.5281/zenodo.14350855
ISSN 2454-9940
www.ijasem.org
Vol 18, Issue 4, 2024

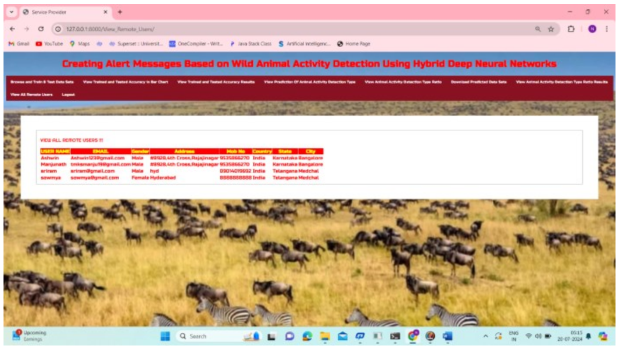

Fig 3: View all Remote Users.

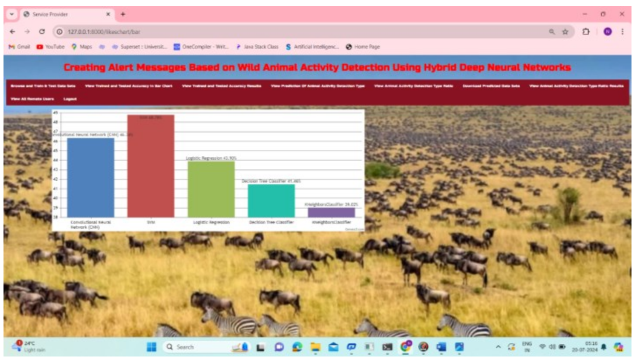

Fig 4: View Accuracy in Bar Chart.

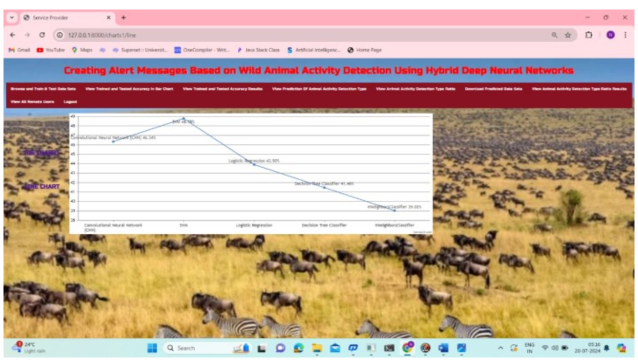

Fig 5: View Accuracy in Pie Chart.

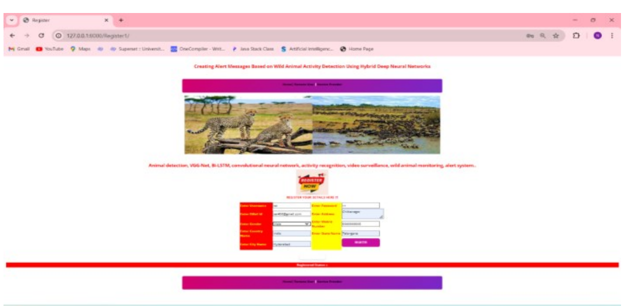

Fig 6: Remote user registration.

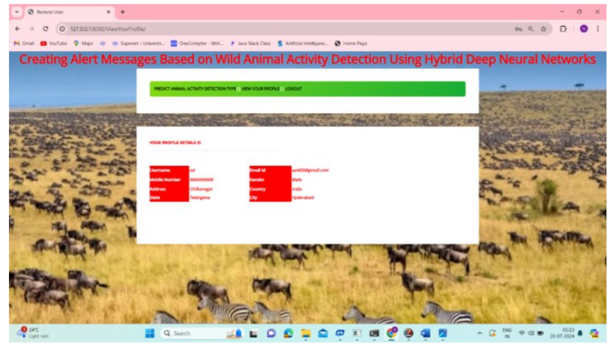

Fig 7: View remote user profile.

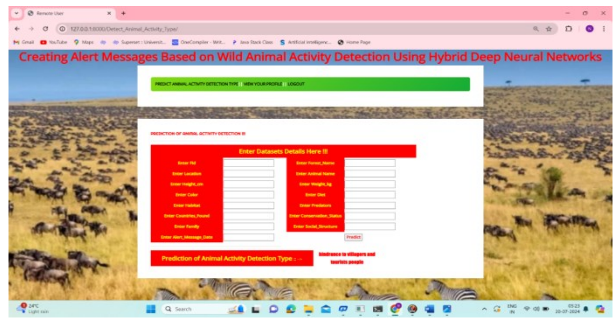

Fig 8: Predicting the Animal Activity Detection.

VI.CONCLUSION

This project introduces the hybrid VGG−19+Bi-LSTMframework for detecting wild animals and helps to monitor the activity of animals. This hybrid approach greatly helps to save the animals from human hunting and humans from animal sudden attacks by sending an alert message to the forest officer. This model introduces novel approaches to upgrade the performance of deep learning techniques in wider applications and real time cases. The proposed model has been evaluated on four different benchmark datasets that contain animal-based datasets—camera trap dataset, wild animal dataset, hoofed animal

INTERNATIONAL JOURNAL OF APPLIED SCIENCE ENGINEERING AND MANAGEMENT
https://doi.org/10.5281/zenodo.14350855
ISSN 2454-9940
www.ijasem.org
Vol 18, Issue 4, 2024

dataset, and CD net data set. The experimental results show the improved performance of our model over various quality metrics. The proposed hybrid VGG−19+Bi-LSTM model achieves above 98% average classification accuracy results and 77.2% mean Average Precision (MAP) and 170 FPS values. Henceforth, the proposed hybrid VGG−19+Bi-LSTM model out performs earlier approaches and produces greater results with lower computation time.

FUTURE ENHANCEMENTS:

In future work will involve increasing the performance of Hybrid Deep Neural Networks in the wild animal activity detection across multiple categories. As the severity of the technology is supposed to be changing with time, so the Hybrid Deep Neural Networks should be improved to enhance classification and detection of the wild animal activity across various areas.

REFERENCES:

[1] S. Aarthi and S. Chitrakala, ‘‘Scene understanding—A survey,’’ in Proc. Int. Conf. Comput., Commun. Signal Process. (ICCCSP), Jan. 2017, pp. 1–4.

[2] N. Ahuja and S. Todorovic, ‘‘Connected segmentation tree—A joint representation of region layout and hierarchy,’’ in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., Jun. 2008, pp. 1–8. [3] T. A. Assegie, P. K. Rangarajan, N. K. Kumar, and D. Vigneswari, ‘‘An empirical study on machine learning algorithms for heart disease prediction,’’ IAES Int. J. Artif. Intell. (IJ-AI), vol. 11, no. 3, p. 1066, Sep. 2022.

[4] N. Banupriya, S. Saranya, R. Swaminathan, S. Harikumar, and S. Palanisamy, ‘‘Animal detection using deep learning algorithm,’’ J. Crit. Rev., vol. 7, no. 1, pp. 434–439, 2020.

[5] M. Cheng, Z. Zhang, W. Lin, and P. Torr, ‘‘BING: Binarized normed gradients for objectness estimation at 300fps,’’ in Proc. IEEE Conf. Comput.Vis. Pattern Recognit., Jun. 2014, pp. 3286–3293.

[6] J. Dai, Y. Li, K. He, and J. Sun, ‘‘R-FCN: Object detection via region-based fully convolutional networks,’’ 2016, arXiv:1605.06409.

[7] M. De Gregorio and M. Giordano, ‘‘Change detection with weightless neural networks,’’ in Proc. IEEE Conf. Comput. Vis. Pattern Recognit.Workshops, Jun. 2014, pp. 409–413.

[8] B. Natarajan, E. Rajalakshmi, R. Elakkiya, K. Kotecha, A. Abraham,L. A. Gabralla, and V. Subramaniyaswamy, ‘‘Development of an end-toend deep learning framework for sign language recognition, translation, and video generation,’’ IEEE Access, vol. 10, pp. 104358–104374, 2022.

[9] W. Dong, P. Roy, C. Peng, and V. Isler, ‘‘Ellipse R-CNN: Learning to infer elliptical object from clustering and occlusion,’’ IEEE Trans. Image Process., vol. 30, pp. 2193–2206, 2021.

[10] R. Elakkiya, P. Vijayakumar, and M. Karuppiah, ‘‘COVID_SCREENET: COVID-19 screening in chest radiography images using deep transfer stacking,’’ Inf. Syst. Frontiers, vol. 23, pp. 1369–1383, Mar. 2021.

INTERNATIONAL JOURNAL OF APPLIED SCIENCE ENGINEERING AND MANAGEMENT
ISSN 2454-9940
www.ijasem.org
Vol 18, Issue 4, 2024
https://doi.org/10.5281/zenodo.14350855

[11] R. Elakkiya, V. Subramaniyaswamy, V. Vijayakumar, and A. Mahanti, ''Cervical cancer diagnostics healthcare system using hybrid object detection adversarial networks,'' IEEE J. Biomed. Health Informat., vol. 26, no. 4, pp. 1464–1471, Apr. 2022.

[12] R. Elakkiya, K. S. S. Teja, L. Jegatha Deborah, C. Bisogni, and C. Medaglia, ''Imaging based cervical cancer diagnostics using small object detection—Generative adversarial networks,'' Multimedia Tools Appl., vol. 81, pp. 1–17, Feb. 2022.

[13] A. Elgammal, D. Harwood, and L. Davis, ''Non-parametric model for background subtraction,'' in Computer Vision—ECCV. Dublin, Ireland: Springer, Jun. 2000, pp. 751–767.

[14] D. Erhan, C. Szegedy, A. Toshev, and D. Anguelov, ''Scalable object detection using deep neural networks,'' in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., Jun. 2014, pp. 2155–2162.

[15] G. Farnebäck, ''Two-frame motion estimation based on polynomial expansion,'' in Proc. 13th Scandin. Conf. (SCIA). Halmstad, Sweden: Springer, Jul. 2003, pp. 363–370.

[16] R. Girshick, ''Fast R-CNN,'' in Proc. IEEE Int. Conf. Comput. Vis. (ICCV), Dec. 2015, pp. 1440–1448.

[17] R. Girshick, J. Donahue, T. Darrell, and J. Malik, ''Rich feature hierarchies for accurate object detection and semantic segmentation,'' in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., Jun. 2014, pp. 580–587.

[18] N. Goyette, P. Jodoin, F. Porikli, J. Konrad, and P. Ishwar, ''Changedetection. Net: A new change detection benchmark dataset,'' in Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. Workshops, Jun. 2012, pp. 1–8.

[19] K. He, X. Zhang, S. Ren, and J. Sun, ''Spatial pyramid pooling in deep convolutional networks for visual recognition,'' IEEE Trans. Pattern Anal. Mach. Intell., vol. 37, no. 9, pp. 1904–1916, Sep. 2015.

[20] J. Imran and B. Raman, ''Evaluating fusion of RGB-D and inertial sensors for multimodal human action recognition,'' J. Ambient Intell. Humanized Comput., vol. 11, no. 1, pp. 189–208, Jan. 2020.

[21] F. Kahl, R. Hartley, and V. Hilsenstein, ''Novelty detection in image sequences with dynamic background,'' in Statistical Methods in Video Processing, Prague, Czech Republic: Springer, May 2004, pp. 117–128.

[22] T. Liang, H. Bao,W. Pan, and F. Pan, ''Traffic sign detection via improved sparse R-CNN for autonomous vehicles,'' J. Adv. Transp., vol. 2022, pp. 1–16, Mar. 2022.

[23] T. Liang, H. Bao, W. Pan, X. Fan, and H. Li, ''Detect Former: Category assisted transformer for traffic scene object detection,'' Sensors, vol. 22, no. 13, p. 4833, Jun. 2022.

[24] G. Li and Y. Yu, ''Deep contrast learning for salient object detection,''in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), Jun. 2016, pp. 478–487.

[25] W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, C.-Y. Fu, and A. C. Berg, ''SSD: Single shot multibox detector,'' in Computer Vision— ECCV. Amsterdam, The Netherlands: Springer, 2016, pp. 21–37.