
ISSN: 2321-2152
IJMECE
International Journal of modern
electronics and communication engineering
E-Mail
editor.ijmece@gmail.com
editor@ijmece.com
www.ijmece.com

ISSN2321-2152
www.ijmece .com
Vol 12, Issue.2 April 2024

CREATING ALERT MESSAGES BASED ON WILD ANIMAL ACTIVITY DETECTION USING DEEP NEURAL NETWORKS

Lavudiya Omkar Nayak [1], Munigati Nitheesh Reddy [2], Thota Laxmi Ganesh [3], Erra Sruchil[4], Mr. K. Radhakrishna[5]
[1,2,3,4] UG Student, Department of ECE,
[5] Assistant Professor, Dept. of ECE
CMR Institute of Technology, 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 from three different benchmark 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.

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 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. These systems offer 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. Research on this work, investigates the advancements in video analysis techniques and complex neural network-based architectures. Recent developments in Deep Learning techniques have produced impressive results in image recognition, classification, and generation tasks [1].

LITERATURE REVIEW

‘‘Scene understanding—A survey,’’

As a prerequisite for autonomous driving, scene understanding has attracted extensive research. With the rise of the convolutional neural network (CNN)-based deep learning technique, research on scene understanding has achieved significant progress. This paper aims to provide a comprehensive survey of deep learning-based approaches for scene understanding in autonomous driving. We categorize these works into four work streams, including object detection, full scene semantic segmentation, instance segmentation, and lane line segmentation. We discuss and analyze these works according to their characteristics, advantages and disadvantages, and basic frameworks. We also summarize the benchmark datasets and evaluation criteria used in the research community and make a performance comparison of some of the latest works. Lastly, we summarize the review work and provide a discussion on the future challenges of the research domain.

‘‘Connected segmentation tree—A joint representation of region layout and hierarchy,’’

This paper proposes a new object representation, called Connected Segmentation Tree (CST), which captures canonical characteristics of the object in terms of the photometric, geometric, and spatial adjacency and containment

properties of its constituent image regions. CST is obtained by augmenting the object's segmentation tree (ST) with inter-region neighbor links, in addition to their recursive embedding structure already present in ST. This makes CST a hierarchy of region adjacency graphs. A region's neighbors are computed using an extension to regions of the Voronoi diagram for point patterns. Unsupervised learning of the CST model of a category is formulated as matching the CST graph representations of unlabeled training images, and fusing their maximally matching subgraphs. A new learning algorithm is proposed that optimizes the model structure by simultaneously searching for both the most salient nodes (regions) and the most salient edges (containment and neighbor relationships of regions) across the image graphs. Matching of the category model to the CST of a new image results in simultaneous detection, segmentation and recognition of all occurrences of the category, and a semantic explanation of these results.

''A Bayesian computer vision system for modeling human interactions,''

We describe a real time computer vision and machine learn ing system for modeling and recognizing human behaviors in a visual surveillance task The system is particularly concerned with detecting when interactions between people occur and classifying the type of in teraction Examples of interesting interaction behaviors include following another person altering one s path to meet another and so forth Our system combines top down with bottom up information in a closed feedback loop with both components employing a statistical Bayesian approach We propose and compare two different state based learning architectures namely HMMs and CHMMs for modeling behaviors and interactions The CHMM model is shown to work much more eciently and accurately Finally to deal with the problem of limited training data a synthetic Alife style training system is used to develop exible prior models for recognizing human interactions We demonstrate the ability to use these a priori models to accurately classify real human behaviors and interac tions with no additional tuning or training.

''Learning and transferring mid-level image representations using convolutional neural networks,''

Convolutional neural networks (CNN) have recently shown outstanding image classification performance in the largescale visual recognition challenge (ILSVRC2012). The success of CNNs is attributed to their ability to learn rich

midlevel image representations as opposed to hand-designed low-level features used in other image classification methods. Learning CNNs, however, amounts to estimating millions of parameters and requires a very large number of annotated image samples. This property currently prevents application of CNNs to problems with limited training data. In this work we show how image representations learned with CNNs on large-scale annotated datasets can be efficiently transferred to other visual recognition tasks with limited amount of training data. We design a method to reuse layers trained on the ImageNet dataset to compute mid-level image representation for images in the PASCAL VOC dataset. We show that despite differences in image statistics and tasks in the two datasets, the transferred representation leads to significantly improved results for object and action classification, outperforming the current state of the art on Pascal VOC 2007 and 2012 datasets. We also show promising results for object and action localization.

‘‘Joint deep learning for pedestrian detection,’’

Feature extraction, deformation handling, occlusion handling, and classication are four important components in pedestrian detection. Existing methods learn or design these components either individually or sequentially. The interaction among these components is not yet well explored. This paper proposes that they should be jointly learned in order to maximize their strengths through cooperation. We formulate these four components into a joint deep learning framework and propose a new deep network architecture1. By establishing automatic, mutual interaction among components, the deep model achieves a 9% reduction in the average miss rate compared with the current best-performing pedestrian detection approaches on the largest Caltech benchmark dataset. Pedestrian detection is a key technology in automotive safety, robotics, and intelligent video surveillance. It has attracted a great deal of research interest [2, 5, 12, 47, 8]. The main challenges of this task are caused by the intra-class variation of pedestrians in clothing, lighting, backgrounds, articulation, and occlusion. In order to handle these challenges, a group of interdependent components are important. First, features should capture the most discriminative information of pedestrians. Well-known features such as Haar-like features [49], SIFT [29], and HOG [5] are designed to be robust to intraclass variation while remain sensitive to inter-class variation. Second, deformation models should handle the articulation of human

parts such as torso, head, and legs. The state-of-the-art deformable part-based model in [17] allows human parts to articulate with constraint. Third, occlusion handling approaches [13, 51, 19] seek to identify the occluded regions and avoid their use when determining the existence of a pedestrian in a window.

‘‘You only look once: Unified, real-time object detection,’’ ‘‘Faster R-CNN: Towards real-time object detection with region proposal networks,’’

—State-of-the-art object detection networks depend on region proposal algorithms to hypothesize object locations. Advances like SPPnet [1] and Fast R-CNN [2] have reduced the running time of these detection networks, exposing region proposal computation as a bottleneck. In this work, we introduce a Region Proposal Network (RPN) that shares full-image convolutional features with the detection network, thus enabling nearly cost-free region proposals. An RPN is a fully convolutional network that simultaneously predicts object bounds and objectness scores at each position. The RPN is trained end-to-end to generate high-quality region proposals, which are used by Fast R-CNN for detection. We further merge RPN and Fast R-CNN into a single network by sharing their convolutional features—using the recently popular terminology of neural networks with “attention” mechanisms, the RPN component tells the unified network where to look. For the very deep VGG-16 model [3], our detection system has a frame rate of 5fps (including all steps) on a GPU, while achieving state-of-the-art object detection accuracy on PASCAL VOC 2007, 2012, and MS COCO datasets with only 300 proposals per image. In ILSVRC and COCO 2015 competitions, Faster R-CNN and RPN are the foundations of the 1st-place winning entries in several tracks. Code has been made publicly available.

EXISTING SYSTEM

The author Zhang et al. proposed wild animal detection using a multi-level graph cut approach for investigating spatial details and a cross-frame temporal patch verification technique for temporal details. The model analyzes the foreground and background details of the camera trap videos. This approach uses a Camera trap and Change Detection net dataset for segmenting the animal object from natural scenes based on cluttered background videos. Although the model produces a high detection rate, fails to perform well in detecting crucial details like location details, and human interruptions. The author [14] proposed animal detection

using Convolutional Neural Network (CNN), and the author proposed animal detection using Iterative Embedded Graph Cut (IEGC) techniques to form regions over images and DeepCNN features and machine learning classification algorithms [15] for classification purposes. Although these models verify the extracted patches are background or animal, still need improvements in classification performance.

Object Detection using deep learning methods attained new heights in computer vision applications. The detection of objects present in images or videos by using object localization and classification techniques gives higher support in detecting various objects present in an image or video. From the extracted results, we can count the number of objects and their activity. This technique is highly used in video surveillance and security-based applications, tracking objects in hidden boxes, monitoring fraudulent activity in public and crowded areas, traffic monitoring and identification of vehicle theft, vehicle number plate recognition, and Object Character Recognition (OCR) [16].

Disadvantages

• The complexity of data: Most of the existing machine learning models must be able to accurately interpret large and complex datasets to detect Wild Animal Activity Detection.

• Data availability: Most machine learning models require large amounts of data to create accurate predictions. If data is unavailable in sufficient quantities, then model accuracy may suffer.

• Incorrect labeling: The existing machine learning models are only as accurate as the data trained using the input dataset. If the data has been incorrectly labeled, the model cannot make accurate predictions.

Proposed System

The proposed architecture comprises five phases of development steps, which includes data pre-processing, animal detection, VGG−19 pre-trained model-based classification, extracting the prediction results, and sending alert messages. In the data pre-processing phase, 45*k* animal images were collected from different datasets such as camera trap, wild animal, and the hoofed animal dataset. The collected images were rescaled to the size of 224×224 pixels and denoised.

In the second phase, we pass the pre-processed images into YOLOR object detection model [39], which identifies the animal present in an image using bounding boxes as illustrated in Fig. 4. In the third

phase, using hybrid VGG−19+Bi-LSTM model we perform image classification tasks and class label prediction was done and animal details are extracted using LSTM Networks. In the fourth phase, we collect the location information of the animal, and the web server creates a SMS alert and sends it to the forest officers. Finally, remedial action will be taken by the forest officers to save the animals and human lives.

Advantages

1) The proposed Hybrid VGG-19+Bi-LSTM model is built using deep neural networks with fine- tuned hyper parameters to yield greater recognition accuracy results.

2) The proposed model aims to achieve outstanding classification results by incorporating novel hybrid approaches.

3) 'The proposed system offers foresters more accurate prediction performance about animal detection and also supports them with faster alert services via SMS.

MODULES

ADMIN

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.

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 authorizes the users.

Remote User

In this module, there are n numbers of users are present.

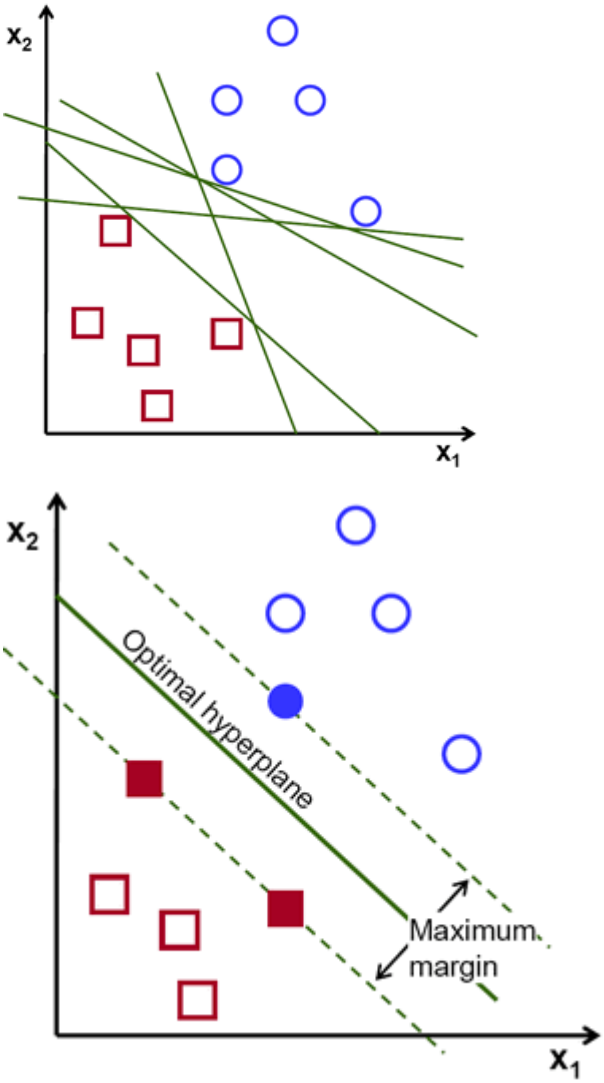
ISSN2321-2152
www.ijmece .com
Vol 12, Issue.2 April 2024

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.

ALGORITHMS

SUPPORT VECTOR MACHINE?

The objective of the support vector machine algorithm is to find a hyperplane in an N-dimensional space(N — the number of features) that distinctly classifies the data points.

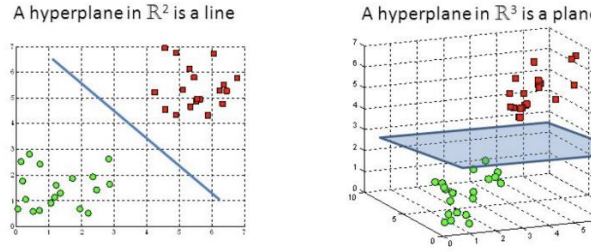


Possible hyperplanes

To separate the two classes of data points, there are many possible hyperplanes that could be chosen. Our objective is to find a plane that has the maximum margin, i.e the maximum distance between data points of both classes. Maximizing the margin distance provides some reinforcement so that future data points can be classified with more confidence.

Hyperplanes and Support Vectors

ISSN2321-2152
www.ijmece .com
Vol 12, Issue.2 April 2024

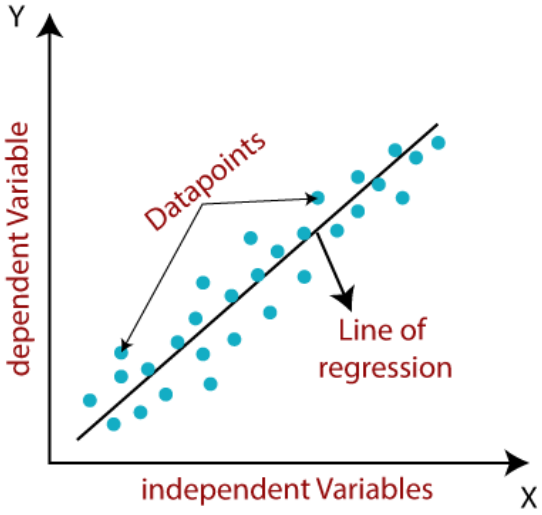


LINEAR REGRESSION

Linear regression is one of the easiest and most popular Machine Learning algorithms. It is a statistical method that is used for predictive analysis. Linear regression makes predictions for continuous/real or numeric variables such as **sales, salary, age, product price,** etc.

Linear regression algorithm shows a linear relationship between a dependent (y) and one or more independent (y) variables, hence called as linear regression. Since linear regression shows the linear relationship, which means it finds how the value of the dependent variable is changing according to the value of the independent variable.

The linear regression model provides a sloped straight line representing the relationship between the variables. Consider the below image:



Mathematically, we can represent a linear regression as:

$$y= a_0+a_1x+ \varepsilon$$

Here,

Y= Dependent Variable (Target Variable)
X= Independent Variable (predictor Variable)
a0= intercept of the line (Gives an additional degree of freedom)
a1 = Linear regression coefficient (scale factor to each input value).
ε = random error

The values for x and y variables are training datasets for Linear Regression model

RANDOM FOREST ALGORITHM

Random Forest is a popular machine learning algorithm that belongs to the

supervised learning technique. It can be used for both Classification and Regression problems in ML. It is based on the concept of **ensemble learning,** which is a process of *combining multiple classifiers to solve a complex problem and to improve the performance of the model.*

As the name suggests, ***"Random Forest is a classifier that contains a number of decision trees on various subsets of the given dataset and takes the average to improve the predictive accuracy of that dataset."*** Instead of relying on one decision tree, the random forest takes the prediction from each tree and based on the majority votes of predictions, and it predicts the final output.

CONCLUSION

This paper introduces the hybrid VGG−19+Bi-LSTM framework 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 anim dataset, hoofed animal dataset, and CDnet dataset. 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 outperforms earlier approaches and produces greater results with lower computation time

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.

[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] Reddy, Kallem Niranjan, and Pappu Venkata Yasoda Jayasree. "Low Power

Strain and Dimension Aware SRAM Cell Design Using a New Tunnel FET and Domino Independent Logic." International Journal of Intelligent Engineering & Systems 11, no. 4 (2018).

[17] Reddy, K. Niranjan, and P. V. Y. Jayasree. "Design of a Dual Doping Less Double Gate Tfet and Its Material Optimization Analysis on a 6t Sram Cells."

[18] Reddy, K. Niranjan, and P. V. Y. Jayasree. "Low power process, voltage, and temperature (PVT) variations aware improved tunnel FET on 6T SRAM cells." Sustainable Computing: Informatics and Systems 21 (2019): 143-153.

[19] Reddy, K. Niranjan, and P. V. Y. Jayasree. "Survey on improvement of PVT aware variations in tunnel FET on SRAM cells." In 2017 International Conference on Current Trends in Computer, Electrical, Electronics and Communication (CTCEEC), pp. 703-705. IEEE, 2017

[20] Karne, R. K. ., & Sreeja, T. K. . (2023). PMLC- Predictions of Mobility and Transmission in a Lane-Based Cluster VANET Validated on Machine Learning. International Journal on Recent and Innovation Trends in Computing and Communication, 11(5s), 477–483. https://doi.org/10.17762/ijritcc.v11i5s.7109

[21] Radha Krishna Karne and Dr. T. K. Sreeja (2022), A Novel Approach for Dynamic Stable Clustering in VANET Using Deep Learning (LSTM) Model. IJEER 10(4), 1092-1098. DOI: 10.37391/IJEER.100454.