International Journal of Recent Technology and Engineering (IJRTE)
ISSN: 2277-3878, Volume-8 Issue-4, November 2019

Deep Learning Methods for monitoring, detecting and measuring Deer Movements for Wildlife Conservation

Neelam Rawat, J.S. Sodhi, Rajesh K. Tyagi

Abstract*: Wildlife Researchers examine and dig video corpus for behavioral studies of free-ranging animals, which included monitoring, analyzing, classifying & detecting, managing, counting etc. Unfortunately, automated visual implementation for challenging real-time scenarios of wildlife is not an easy task especially for classification and recognition of wildlife-animals and estimate the sizes of wildlife populations. The aim of this paper is to bring state-of-the-art results from raw sensor data for learning features advancing automatic implementation and interpreting of animal movements from different perspectives. Also, turnout with an objectness score from object proposals generated by Region Proposal Network (RPN). The imagery data are captured from the motion sensor cameras and then through RCNN, Fast RCNN and Faster RCNN, it automatically are segmented and recognized the object with its objectness score. ConvNet automatically process these images and correctly recognizing the object. Experimentation results demonstrated prominent deer images with 96% accuracy with identifying three basic activities sleeping, grazing and resting. In addition, a measured implementation has been shown among CNN, RCNN, Fast RCNN and Faster RCNN.

***Keywords*: Wildlife, Ecosystem, Ecology, Deep Learning, Convolutional Neural Network, Region Proposal Network, motion sensor cameras**

I. INTRODUCTION

Understanding wildlife conservation is the practice of protecting animal species and balancing the ecosystem so that to provide the sustenance in wildlife preservation and management. Population of living species becoming extinct due to climate change, unregulated hunting and poaching, over-exploitation, deforestation etc. Despite these issues, need of wildlife conservation in India is now questioned. Apparently, several wild animal species, forest brimming towards destruction and disappearance. To have a finer knowledge of wildlife ecosystem, more convincing and virtuous implementation is required for wildlife species [27]. For this, an accurate method for auto-implementation of animal identification, detection and monitoring from camera-trap images should be contrived [28][29]. Effective surveillance is crucial for every detection and successful monitoring for terrestrial animals and their habitats.

From camera-trap images [1], processing of extensive amount of images and videos are very expensive and time consuming which may further be a major obstacle for research scientists and ecologists. In this paper, a legal

framework is proposed to frame an auto-implementation animal detection and recognition in the wild. In wildlife monitoring, the standard data used is population-ratio (Figure 2). Statistically, small samples are not reliable, but may be the animals having same habitat and living, in that case, a small sample of animals can represent the population. Monitoring the frequency relying on individual animals in the detection and identification process. Estimating the density of animals without use of individual recognition may also speed-up the underlying process [5][6].

In this paper, a legal framework is proposed to frame an auto-implementation animal detection and recognition in the wild. In wildlife monitoring, the standard data used is population-ratio. Statistically, small samples are not reliable, but may be the animals having same habitat and living, in that case, a small sample of animals can represent the population. Monitoring the frequency relying on individual animals in the detection and identification process [2][3]. Estimating the density (or counting) of animals without use of individual recognition may also speed-up the underlying process [10][26].

From recent years, Deep Learning (DL) techniques proves to be boom in the area of computer vision [17]. The objective of this paper is to do measured implementation on DL methodologies and through these, monitoring and detecting the deer movements and activities. Although, it may frame for any wildlife species, but this multi-task learning method make a clear picture of the frame that able to identify effectively and transit that in model [4][16].

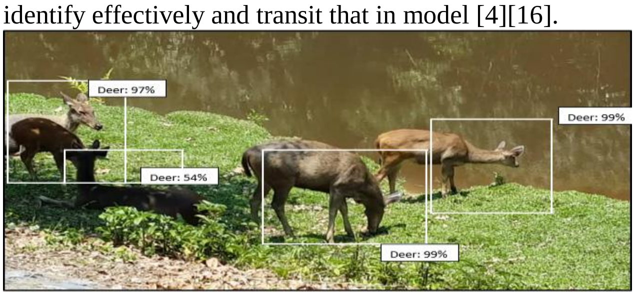

Figure 1: Recognizing Deer with Objectness Scores

Revised Manuscript Received on November 15, 2019

Neelam Rawat*, Research Scholar, Amity University, Noida, India. Email: neemarawat11@gmail.com

J. S. Sodhi, Senior Vice President, Amity University. Noida, India Email: jssodhi@akcgroup.com

Rajesh K. Tyagi, Head & Professor, Computer Science, Amity University, Gurugram, India. Email: profrajeshkumartyagi@gmail.com

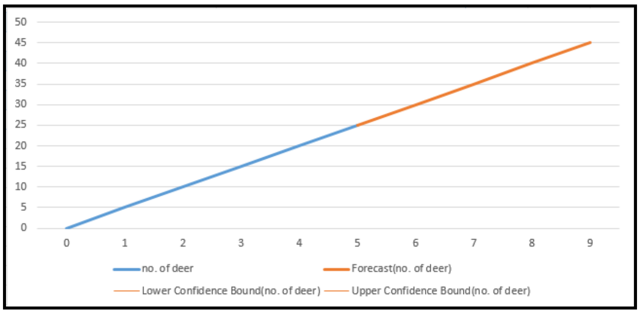

Figure 2: Population Ratio of Deer with Confidence level of detection accordingly.

A. CNN, RCNN, Fast RCNN and Faster RCNN

CNN (Convolutional Neural Network) is a class of feed forward artificial neural networks which is designed as a multi- layer perceptron (deep learning algorithm) to identify different objects, recognize each object individually with minimal processing [7][8]. CNN applications are image recognition, text analytics, for sound and graph data etc. CNN not able to perceive images, rather consider images as volumes with its width and height measured by its pixels. For each pixel, some numbers define intensity of R, G, and B and that identified number will be a component among three. These numbers are initially raw features, which are feed in CNN to find the significance of those numbers as objects. This filtering process continues until it identifies the patterns in the pixels [5][6]. The major shortcomings of using CNN are if features are homogeneous then no issues in processing and identifying images but will be slow for different scales of heterogeneous features and ineffective also. CNN is cost effective if have had a large dataset with good GPU because the computational cost of CNN is quite high as works well with high dimensional data.

RCNN (regions with CNN features) or Region-based Convolutional Neural Network that apply to deep models for object recognition [9][10]. RCNN defines the region from any image by using a bounding box so that the focus on a single region at a time to avoid interference. In RCNN, region is defined with selective search algorithm with continual of resizing it so that of equal region sizes before feeding to CNN for classification process (Figure 3). Selective Search selects multiple regions

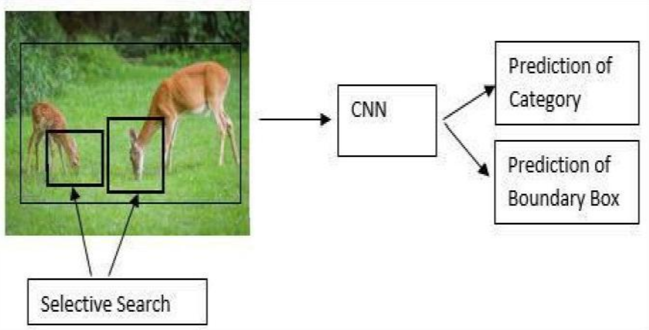

Figure 3: RCNN with defined images and selective search algorithm for classification

RCNN models effectively extract target images from pre-trained CNNs but the computational speed becomes slow as thousands of regions can be formed from a single image which require thousands of computation perform to suitable for real-time applications frequently based on number of regions formed [11][12].

Fast RCNN (Fast Region-based Convolutional Neural Network) extracts independent features from each region that improves the performance compare to RCNN and simple CNN because through Fast RCNN computation is now only on an image as a whole [18]. This time, no need to feed thousands of regions to CNN every time but operation is done only once per image to generate feature. (Figure 4). In addition, Region of Interest (RoI) pooling layer will defined directly the shape of each region with height and width specifications that makes RoI layer to extract features of same shapes from RoI of different shapes [19].

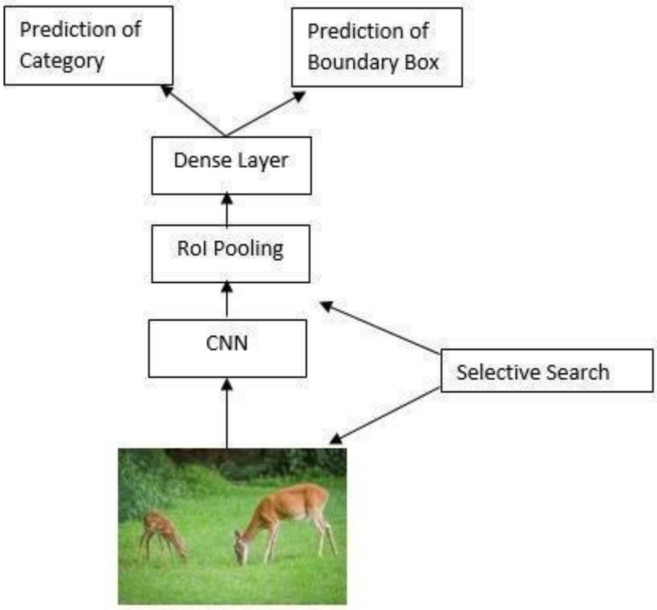

Figure 4: Fast RCNN Model with RoI Pooling and Selective Search algorithm for classification

Fast RCNN model achieves excellent object detection accuracy compare to the previous defined models. However, the computational and training powers in terms of space and time is very expensive and the process of detecting objects is quite slow [13].

Faster RCNN (Faster Region-based Convolutional Neural Network) defines the precise detection of objects by replacing selective search to region proposal network (RPN). Through RPN, number of proposed generated regions will be reduced and ensure the object detention process in precise manner [20][21]. Faster RCNN generates high-quality regions from selective search algorithms to RPN including the detection of boundary boxes and category with binary category prediction and bounding box for anchor box in RPN (for predicting whether there is an object or not with defined bounding box for objects). (Figure 5). Therefore, faster RCNN has two main modules: deep fast RCNN for the use of detected regions and CNN for region, which defines a unified model for object detection and identification process [14].

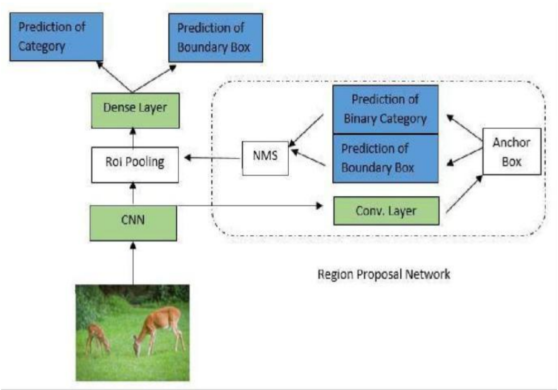

Figure 5: Faster RCNN Model with Region Proposal Network (RPN) for classification

Further, RoI pooling layer replaced by RoI alignment layer to predict pixel-level positions of objects that is not done in Faster RCNN. So we do have Mask RCNN which modifies the faster RCNN model by using RoI alignment layer for spatial information of an object [15][22]. This pixel-level semantic detection could be a challenge and limitation for faster RCNN.

II. MATERIAL AND METHODS

CNN (Convolutional Neural Network), the simplest deep learning methods for widely used in detecting images. CNN considers as a part of artificial neural networks for identifying different objects and activities with recognizing each object separately of minimal computational processing. CNN is not bound only for image processing, but its efficacy proves to work on medical diagnosis, business pattern detections and so on. Let's consider firstly the inner working of CNN (Figure 6) as:

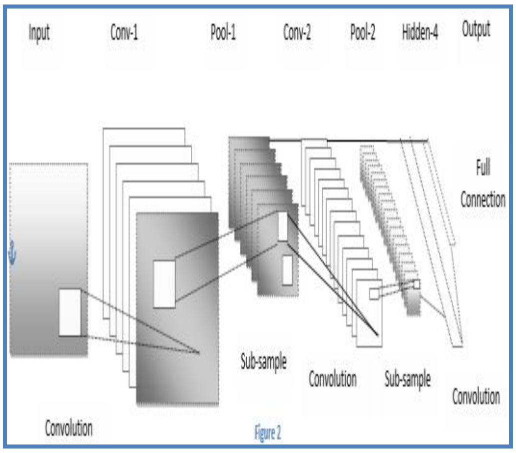

Figure 6: CNN Framework

Step 1: Consider an image as an input (Figure 7. below).

Figure 7: Input Image

Step 2: Various regions are formed by splitting the input image (Figure 8. Below).

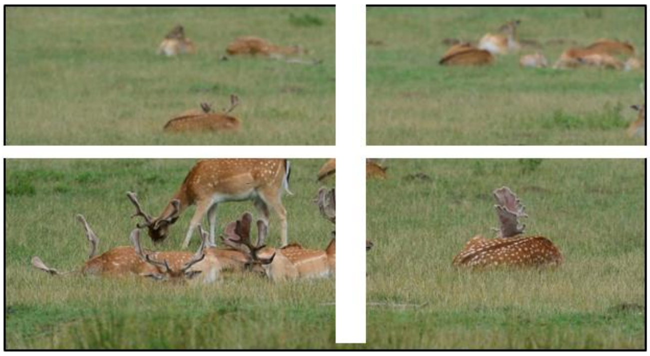

Figure 8: Split of Input Image

Step 3: We marked individual image from each region
Step 4: Propel all these regions (images) to CNN and arrange it into various classes.
Step 5: Finally, we detect the object and merge all the regions to get the original image (Figure 9. Below).

Figure 9: Bounding Box generated for each Region

Worriment is the objects in images may have several aspect ratios and spatial locations [23][24]. Such as, in particular, cases, the object might overlay most of the image, but in others object might have a very small percentage of image. In given Figure 6. on left hand side of the image, object covering most of the images compare to other side of the same. In addition to, shape and size of objects may also vary. With the consideration of all these factors, we would require voluminous regions that may result with excessive computational time [25].

Therefore, we need to scale down the number of regions and use region-based CNN that screened the regions in particular. To avoid such issues, we need to focus on single regions at a time. In RCNN (Region-based Convolutional Neural Network), CNN is coerced to concentrate on a particular region to recede the interferences. RCNN uses Selective Search Algorithms i.e., an object proposal algorithm that uses the indications like textures, intensity, color etc. to predict the possible locations of objects. These generated boxes then go as an input to CNN classifier with the following important aspects:

- Run the algorithm to find all possible objects
- These regions further get feed to CNN so that particular class get detected by SVM
- Optimize the regions for identifying the patterns in the images.

Following are the steps to describe how selective search (SS) for RCNN works:

Step 1: Initially, consider an image as an input.
Step 2: To have multiple regions, segment the images.
Step 3: The technique commingles related regions to form a particular region based on texture, color, size and shape similarity.
Step 4: At last, final location of object is being produced from these regions (i.e., Region of Interest)

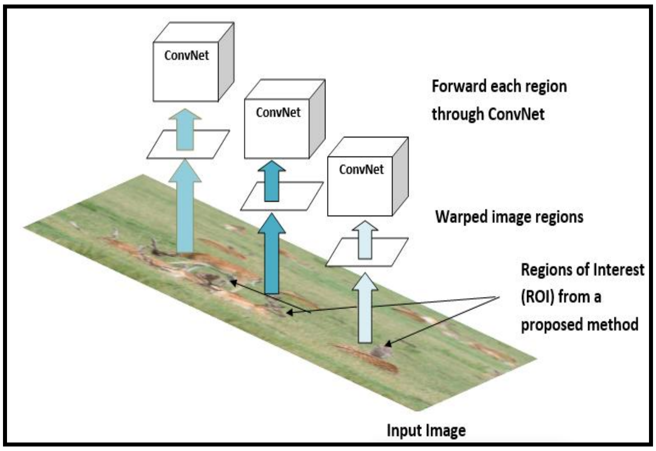

Figure 10: RCNN Framework

Succinct steps are followed in RCNN to detect objects:

- Take a consequent CNN network for the classification of image. We have N classes for this purpose.
- Through selective search, ROI is proposed and that region have target objects with varying sizes.
- From each image's ROI, match the input size of CNN by reshaping all regions.
- From these resultant regions, we classify objects and background by using SVM as one binary SVM for each class.
- To have final bounding boxes, we train linear regression model for each object which get identified from images.

In the above Figure 9, a bounding box is generated for each region to predict the targeted object. RCNN used region proposals at pixel levels into consequent CNN for detection purpose. Therefore, both the methods of RPN network through Selective Search (SS) algorithm and detection network get uncoupled, through, which if error occurred in SS network then it, may affect the detection network directly. So it is preferred that SS and detection network correlate with each other. However, training with RCNN is very expensive and slow as well because if we consider N images and 1000 features for extraction from them by using CNN, then it needs N*1000 CNN features which may make RCNN very slow. As around 50-60 seconds of time will take to have a prediction of individual images that make it a cumbersome model and practically not possible to host.

We can have all ROI by executing once for each image through Fast RCNN (Fast Region-based Convolutional Neural Network) because Fast RCNN uses RPN at feature map level instead pixel level. In Fast RCNN, we consider an image as an input (Figure 7.) which passed through CNN for featuring images to get ROI of each. The process use ROI pooling layer to get it shaped again into fixed sizes so that to correlated with Fully-Connected (FC) layers.

Following are the steps taken for using Fast RCNN:
Step 1: Consider an image input.
Step 2: Pass this image to CNN for generating Region of Interest (RoI)
Step 3: Apply ROI pooling layer tall the regions to reshape this image, then pass it to fully connected network.
Step 4: On the top of fully connected network, Softmax (Linear) layer is used to generate classes as output. In parallel, a regression layer is used for bounding box coordinates to predict the classes.

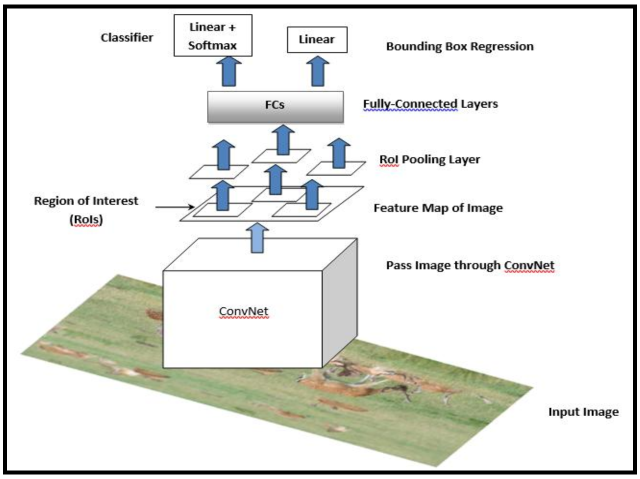

Figure 11: Fast RCNN Framework

From the above Figure 11, it confirms that Fast RCNN resolute the issues of RCNN, as it passes single image instead to CNN and uses single model also for generating the boundary boxes for classification and extraction of features. Although, Fast RCNN uses SS algorithm for finding RoIs that makes the process very slow and time consuming. Although Fast RCNN detects an image in and around 2-3 seconds that may result had better compare to RCNN, yet it will not be effective for large real-time datasets. For such datasets, we do have another option of Faster RCNN (Faster Region-based Convolutional Neural Network) i.e., a modified version of Fast RCNN. In contrast, Faster RCNN (Figure 12) uses that considers image feature map as an input to produce the object proposals with objectness score as output (for reference, mentioned in Figure 1 as an object with objectness score).

Following are the steps taken for using Faster RCNN:
Step 1: Firstly, we take an input of image that returns a feature map while passing from CNN.
Step 2: For this Image Feature Map, we apply Region Proposal Network to generate the objectness score with its object proposal.
Step 3: To bring down the similar image, apply RoI pooling layer to all the regions and pass it to FC network.
Step 4: For class prediction, use the coordinates of boundary boxes by using linear regression layer.

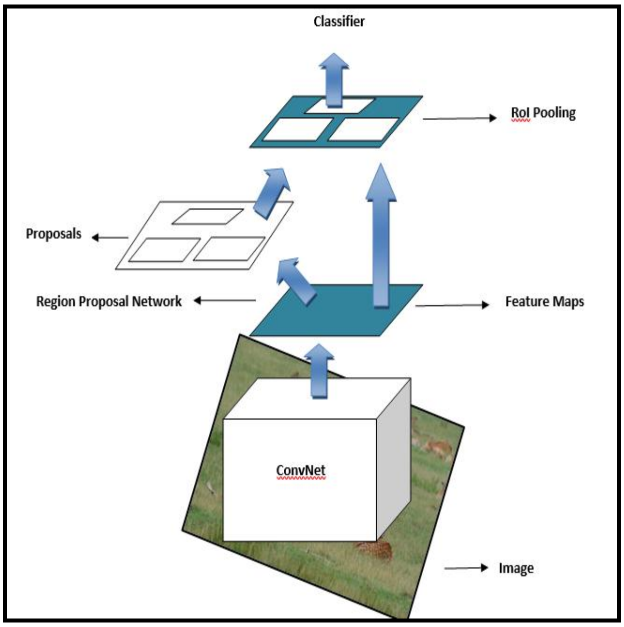

Figure 12: Faster RCNN Framework

III. RESULT AND DISCUSSION

By using camera-trap images, this practice focus only on deers without considering the images having more than one species. With this methodology, we classify and label the activities of captured deer. As given in Figure 13. (from left onwards), capture-1 represents sleeping activity, capture-2 represents grazing activity and capture-3 represents resting activity.

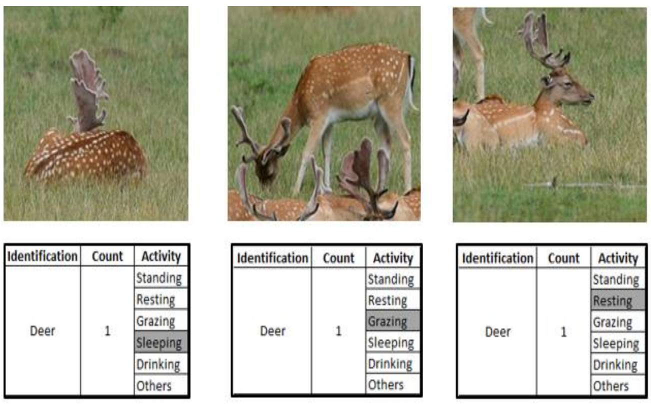

Figure 13: Recognition of Deer Movements

IV. DISCUSSION

Our analytical method allowed us to make a clear and measured results of deer movements and on that basis only we may able to make a summary of Deep Learning methods below (Table 1.):

Method	Prominence	Prophecy time / image	Restrictions
CNN	Arrange each regions into different classes by splitting it into regions.	–	Very high computation time as it requires numerous regions for accurate predictions.
RCNN	Selects only bunch of regions, if find object the perform image extraction on or above 1500 regions from each image.	50-60 seconds	Since regions are passed separately from CNN, so have high computation time. In addition, it is very expensive and slow.
Fast RCNN	Images are passed only once through CNN then extracts featured maps. For predictions, it passes only one image to CNN by using only 1 model instead, 3 and then feature maps are extracted.	2-3 seconds	Computation time is quite high but selection of search make it slow.
Faster RCNN	Use region proposal network (RPN) instead some selective search.	0.19 seconds	Although, proposal of object takes time yet RPN makes it very faster

Table 1: Measured Implementation of Deep Learning Methods Implications

V. IMPLICATIONS

Our methodology necessitates minimum training set of 15-20 images which shorten the efforts of building a training set and code this perspective method is more convenient for wildlife researchers. Because of various systems working one after another, the further performance depends on how the system performed previously. All in all, encore from neoteric object detection algorithms such as YOLO (You Only Look Once), RatinaNet, SSD (Single Shot Multi-box Detector). These methods and algorithms are based on machine learning, imposing umpteen understanding of mathematical and deep learning frameworks. The work is a novel method for conferring the reliability for monitoring, detection and analyzing deer movements with auto-indexing from wildlife videos (Figure 1). The method handles occlusions, background clutters and most hidden deer behind vegetation. Experiments showed that accurate and robust detection of deer movements.

VI. ACKNOWLEDGMENT

We thank to Amity Institute of Forestry and Wildlife - Amity University for assistance in the field. We specially thanks to all anonymous reviewers for helpful comments on earlier drafts as well.

VII. CONFLICT OF INTEREST

The authors declare that there is no conflict of interest's regarding the publication of this paper.

REFERENCES

1. Norouzzadeh, M.S., Nguyen, A., Kosmala, M., Swanson, A., Palmer, M.S., Packer, C. and Clune, J., 2018. Automatically identifying, counting, and describing wild animals in camera-trap images with deep learning. Proceedings of the National Academy of Sciences, p.201719367.
2. Yazdi, M. and Bouwmans, T., 2018. New trends on moving object detection in video images captured by a moving camera: A survey. Computer Science Review, 28, pp.157-177.

3. Stahl, T., Pintea, S.L. and van Gemert, J.C., 2018. Divide and Count: Generic Object Counting by Image Divisions. IEEE Transactions on Image Processing.
4. Pawar, K. and Attar, V., 2018. Deep learning approaches for video-based anomalous activity detection. World Wide Web, pp.1-31.
5. Verma, G.K. and Gupta, P., 2018. Wild Animal Detection Using Deep Convolutional Neural Network. In Proceedings of 2nd International Conference on Computer Vision & Image Processing (pp. 327-338). Springer, Singapore.
6. Shoemaker, K.T., Heffelfinger, L.J., Jackson, N.J., Blum, M.E., Wasley, T. and Stewart, K.M., 2018. A machine-learning approach for extending classical wildlife resource selection analyses. Ecology and evolution, 8(6), pp.3556-3569.
7. Nguyen, H., Maclagan, S.J., Nguyen, T.D., Nguyen, T., Flemons, P., Andrews, K., Ritchie, E.G. and Phung, D., 2017, October. Animal recognition and identification with deep convolutional neural networks for automated wildlife monitoring. In Data Science and Advanced Analytics (DSAA), 2017 IEEE International Conference on (pp. 40-49). IEEE.
8. Zhou, X., Gong, W., Fu, W. and Du, F., 2017, May. Application of deep learning in object detection. In Computer and Information Science (ICIS), 2017 IEEE/ACIS 16th International Conference on (pp. 631-634). IEEE.
9. Tian, B., Li, L., Qu, Y. and Yan, L., 2017, August. Video Object Detection for Tractability with Deep Learning Method. In Advanced Cloud and Big Data (CBD), 2017 Fifth International Conference on (pp. 397-401). IEEE.
10. Chattopadhyay, P., Vedantam, R., Selvaraju, R.R., Batra, D. and Parikh, D., 2017, July. Counting Everyday Objects in Everyday Scenes. In CVPR (pp. 4428-4437).
11. Valletta, J.J., Torney, C., Kings, M., Thornton, A. and Madden, J., 2017. Applications of machine learning in animal behaviour studies. Animal Behaviour, 124, pp.203-220.
12. Gomez, A., Salazar, A. and Vargas, F., 2016. Towards automatic wild animal monitoring: Identification of animal species in camera-trap images using very deep convolutional neural networks. arXiv preprint arXiv:1603.06169.
13. He, K., Zhang, X., Ren, S. and Sun, J., 2016. Deep residual learning for image recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 770-778).
14. Redmon, J., Divvala, S., Girshick, R. and Farhadi, A., 2016. You only look once: Unified, real-time object detection. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 779-788).
15. Wang, X., 2016. Deep learning in object recognition, detection, and segmentation. Foundations and Trends® in Signal Processing, 8(4), pp.217-382.
16. Agarwal, V.K., Sivakumaran, N. and Naidu, V.P.S., 2016. Six Object Tracking Algorithms: A Comparative Study. Indian Journal of Science and Technology, 9(30).
17. Szegedy, C., Erhan, D. and Toshev, A.T., Google Inc, 2016. Object detection using deep neural networks. U.S. Patent 9,275,308.
18. Arteta, C., Lempitsky, V. and Zisserman, A., 2016, October. Counting in the wild. In European conference on computer vision (pp. 483-498). Springer, Cham.
19. Stern, U., He, R. and Yang, C.H., 2015. Analyzing animal behavior via classifying each video frame using convolutional neural networks. Scientific reports, 5, p.14351.
20. Cai, Z., Fan, Q., Feris, R.S. and Vasconcelos, N., 2016, October. A unified multi-scale deep convolutional neural network for fast object detection. In European Conference on Computer Vision (pp. 354-370). Springer, Cham.
21. AlpaydÕn, G., An Adaptive Deep Neural Network for Detection, Recognition of Objects with Long Range Auto Surveillance.
22. Chen, G., Han, T.X., He, Z., Kays, R. and Forrester, T., 2014, October. Deep convolutional neural network based species recognition for wild animal monitoring. In Image Processing (ICIP), 2014 IEEE International Conference on (pp. 858-862). IEEE.
23. Sharif Razavian, A., Azizpour, H., Sullivan, J. and Carlsson, S., 2014. CNN features off-the-shelf: an astounding baseline for recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition workshops (pp. 806-813).
24. Jacoby, D.M., Brooks, E.J., Croft, D.P. and Sims, D.W., 2012. Developing a deeper understanding of animal movements and spatial dynamics through novel application of network analyses. Methods in Ecology and Evolution, 3(3), pp.574-583.
25. Wang, Q., Chen, F., Xu, W. and Yang, M.H., 2011, September. An experimental comparison of online object-tracking algorithms. In Wavelets and Sparsity XIV (Vol. 8138, p. 81381A). International Society for Optics and Photonics.
26. Lancia, R.A., Kendall, W.L., Pollock, K.H. and Nichols, J.D., 2005. Estimating the number of animals in wildlife populations.
27. Rawat, N., Sodhi, J.S. and Tyagi, R., Wild Life Protection by Moving Object Data Mining-Discover with Granular Computing.
28. Rawat, N., Sodhi, J.S. and Tyagi, R.K., 2014, November. An algorithmic approach for analysis of animal movement with granular computing in relation with data mining. In 2014 International Conference on Contemporary Computing and Informatics (IC3I) (pp. 224-229). IEEE.
29. Rawat, Neelam; S. SODHI, J.; K. TYAGI, Rajesh. Analysis and tracking of animal movements through granulation of temporal domain (GTD). International Journal of Engineering & Technology, [S.l.], v. 7, n. 4.5, p. 501-505, sep. 2018. ISSN 2227-524X

AUTHORS PROFILE

Neelam Rawat, Research Scholar from Amity University. Research Interests: Pattern Recognition, Artificial Intelligence, Machine Learning, Deep Learning. She has an experience of 17 years in academics and published papers in reputed journals and conferences. She aspires to motivate my faculties and students towards research and development program, who are the best resources of the Institute. Also, aspire to make work place as a center of excellence in traditional and innovation interdisciplinary, multidisciplinary research and development program.

Dr. J. S. Sodhi CIO and Vice President of Amity Education Group and AKC Group of Companies. He has 25 years of Professional experience with Doctorate in Information Security. He is leading team of 150 IT professionals for overall IT management responsibility of Amity Institutions throughout 35 locations in India & 10 International locations and AKC Group of Companies. is also Executive Director of CCFIS (Cyborg Cyber Forensics and Information Security Pvt Ltd), a leading security research Organization under Amity Innovation Incubator. He has extreme caliber to motivate the human resources of the institute for best productivity by using value based persuasive talk.

Dr. Rajesh K. Tyagi Professor in Computer Science and Engineering at AMITY University Gurgaon. He has committed to introduce the new state of art cutting edge labs facility from various funding agencies such as AICTE/ DST/ CSI/ IEEE/ ACM for getting new patents/ courses/ technologies in the Institute meticulously. He also has caliber to setup a cutting edge value added placements course such as dot net PHP/ Linux/ DBMS/ CCNA/ Android-based programs in work place for enhancement of learning and competitive attitude among students. He has made a link between an industry-academia for industry training as well as good placements.