International Journal of Scientific Research in Science and Technology
Print ISSN: 2395-6011 | Online ISSN: 2395-602X (www.ijsrst.com)
doi : https://doi.org/10.32628/IJSRST

Wild Animal Detection System Using Deep Convolutional Neural Networks

Meenatchi K[1], Thibishini V[1], Vaisnavi K[1], Mrs. R. Ahila[2]

[1]UG Scholar, Department of Computer Science and Engineering, School of Engineering, Avinashilingam Institute for Home Science and Higher Education for Women, Coimbatore, Tamil Nadu, India

[2]Assistant Professor, Department of Computer Science and Engineering, School of Engineering, Avinashilingam Institute for Home Science and Higher Education for Women, Coimbatore, Tamil Nadu, India

ABSTRACT

Article Info
Volume 8, Issue 3
Page Number : 989-996

Publication Issue
May-June-2021

Article History
Accepted : 18 June 2021
Published : 27 June 2021

Animal detecting and monitoring has always been a challenging in research area. Most of the animal detecting and monitoring processes rely on commercial wild camera trap to take wild animal pictures which are triggered by some sort of sensor techniques. The taken images still need human to collect and get analysed with tremendous amount of effort. In a wild environment, the cost for deploying, collecting, analyzing is quite significant. In progress of AI technique, there are mature tools that can be used to analyse the collected images. It can be utilized to solve the wild animal detecting and monitoring problem using Deep Convolutional Neural Networks. The idea is simple to run AI on Raspberry Pi locally to detect a wild animal and then it verifies the images. Then it sends a message through GSM module with no need of internet connection and gives an ultrasonic buzzer sound to divert a wild animal. It is trying to propose an end –to-end solution which could potentially reduce the loss of humans, animals and capitals using animal detecting system using Deep convolutional Neural Networks.

Keywords: Artificial Intelligence, Deep Convolutional Neural Networks, Raspberry PI, Retinex filtering, Softmax.

I. INTRODUCTION

Animal monitoring and analysis an active research field [1] since last many decades. It focuses on animal 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 using a camera-trap database that has candidate animal proposals using multilevel graph cut in the spatio-temporal domain. These proposals are used to create a verification phase that identifies whether a given patch is animal or background. It is designed animal detection model using self-learned

Deep Convolution Neural Network (DCNN) [3] 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. The intensive results show that a detection model using DCNN features provides accuracy of 91.4% on standard camera-trap dataset.

Visual recognition has been gaining popularity in biodiversity preservation and management. Since launching AI for Good initiative, this have been working with biodiversity researchers and practitioners [2] to deliver Animal image recognition machine learning models and tools. In first foray, this area for Wild Detection which aligned with one of the goals to use data science consulting to aid in the preservation and management of the planet's Animal and environment. The goal was to build a model for visual recognition of specific kinds of animals.

II. RELATED WORK

Computer vision techniques [3] are applied to perform automatic wildlife surveying and animal monitoring. Animal detection in aerial videos is challenging because of the complexity of wild environments. The method for moving animal detection is proposed by taking advantage of global patterns of pixel motion. In the video dataset, where animals make obvious movement against the background, motion vectors of each pixel are estimated by applying optical flow methods [4]. A coarse segmentation then removes most parts of the background by applying a pixel velocity threshold. Based on the segmented regions, another threshold was employed to filter out negative candidates that could belong to the background. The pros and cons of this method are discussed.

Wild animal detection helps wildlife researchers to analyze and study wild animal habitat and behaviour. Discriminative Feature-oriented Dictionary Learning (DFDL) [5] was utilized for learning discriminative features of positive images that have animals present in positive class, in addition of negative images that do not have animals present in that class. But, this approach has low performance for detection of visual wild animals. Hence, in this paper, Multi-Cluster Feature Selection (MCFS) is proposed for unsupervised feature selection and wild animal detection. Those features are chosen, which the multi-cluster structure of the data is well preserved. Based on spectral analysis approaches, the proposed method suggests a principled manner for calculating the correlations among various features without label information. Thus, the proposed technique handles the data with multiple cluster structure. The experimental results show that the proposed approach provides the better results.

Efficient and reliable monitoring of wild animals in the natural habitats is essential to inform conservation and management decisions. Auto convert cameras or "camera traps" are being an increasingly popular tool for wild life monitoring due to 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 presence a major obstacle to scientists and ecologists to monitor wild life in an open environment. Leveraging on recent advances in deep learning techniques in computer vision, to build automated animal recognition in the wild, aiming at an automated wild life monitoring system.

Now a days, world has made computers [6] an inseparable part of their life as computers are used for performing the entire work of humans with better accuracy and efficiency. Visual scene analysis is a high-level tasks that acquire knowledge from videos or digital images that comes under the domain of computer vision. Object Detection is a field of computer vision and image processing which involves detecting objects of varying class (animal, humans or cars) present in images and videos. Some well researched applications of object detection are in the domain of car detection, face detection, image retrieval and video surveillance. This survey especially focuses on to examine the different images and videos based object detection methods to support various environments. The main objective of this research is to study about different images and videos based object detection methods used for detecting and solving images and videos based object detection problems. It provides detailed information about the different object detection techniques in various environments. Finally, comparisons are made for different object detection methods used in different images and videos environments.

A novel method for object recognition based on hybrid local descriptors is presented in [7]. This method utilizes a combination of a few approaches (SIFT - Scale-invariant feature transform, SURF - Speed up Robust Features) and consists of second parts. The applicability of the presented hybrid methods are demonstrated on a few images from dataset. Dataset classes represent big animals situated in Slovak country, namely wolf, fox, brown bear, deer and wild boar. The presented method may be also used in other areas of image classification and feature extraction. The experimental results show, that the combination of local descriptors has a positive effect for object recognition.

Animal detection-based study [8] is useful in many real-life applications. Techniques involved in animal detection are useful in observing the locomotive behaviour of the engaged animal and in result it prevent harmful interruption of animals in residential areas. There are some branches of research in animal detection. Some of these branches will therefore be discussed in this journal. Humans have developed many algorithms and techniques to gain a better understanding of animal behaviour. For early preventive measures, these technologies can also serve as a warning system for humans from encroachment of dangerous wild animals. Such tasks can be reduced to three main branches, namely animal detection, tracking and recognition. Through this a new approaches for study and a variety of technologies/algorithms implemented in the past are identified and appropriate ways for solving the research gaps are suggested to fill the gap.

Monitoring animals in the wild without disturbing them is possible using camera trapping framework, which is a technique to study wildlife using automatically triggered cameras [9] and produces great volumes of data. However, camera trapping collects images often result in low image quality and includes a lot of false positives (images without animals), which must be detection before the post processing step. It presents a two-channelled perceiving residual pyramid networks (TPRPN) for camera trap images objection. The TPRPN model [10] attends to generating high-resolution and high-quality results. In order to provide enough local information this extract depth cue from the original images and use two-channelled perceiving model as input for training the networks. Finally, the proposed three-layer residual blocks learn to merge all the information and generate full size detection results. Besides, it constructs a new high-quality dataset with the help of Wildlife Thailand's Community and enamel Organization. Experimental results on dataset demonstrate the method is superior to the existing object detection methods.

Recent studies in computer vision [11] have provided new solutions to real-world problems. This focus on using computer vision methods to assist in the study of kangaroos in the wild. In order to investigate the feasibility, to build a kangaroo image dataset from collected data from several national parks across the State of Queensland. To achieve reasonable detection accuracy, we explored a multipurpose approach and proposed a framework based on the state-of-the-art Deformable Part Model (DPM). Experiments show that the proposed framework outperformed the state-of-the-art methods on the proposed dataset. Also, the proposed vision tools are able to help our field biologists in studying kangaroo related problems such as population tracking for activity analysis.

III. SYSTEM MODEL

3.1 EXISTING MODEL

This is design to build a wild animal pest repellent device with combination of passive infrared (PIR) sensor and ultrasonic signal based on microcontroller as system controller. The PIR sensor is used to detect the presence of wild animal objects and ultrasonic signals to interfere with the hearing. The design of the system is built based on microcontroller as the system controller. The system as a whole includes hardware and software. The design of hardware consists of the system design on the transmitter side and the system design on the receiver side, the software in the system are algorithms using C language programming. Findings – The resulting repellent device can detect animals approaching up to a distance of 5 m and may interfere with its hearing with a 40 kHz ultrasonic frequency up to a distance of 20 m.

Figure 1 Existing Model

The system also uses remote monitoring devices using 433 MHz radio frequency up to a distance of 60 m. Research Limitations/Implications – Each animal has different hearing frequencies, as well as some wild animals, but the hearing frequencies of wild animals are generally at ultrasonic frequencies. The frequency of animal hearing may vary from audio frequency to ultrasonic frequency, so ultrasonic wave emission testing with varying frequencies is required. Practical Implications – This research combines systems on transmitters and receivers, with real-time monitoring of wild animal positions, and it can be possible to monitor the position of more detailed animals by installing more types of sensors as well as increasing the number of sensors.

3.2 PROPOSED MODEL

In this work, it has designed a model that verifies the animal and background patches from the camera-trap images. The challenges associated with the model are the huge variations in background such as dynamic texture of background, change of position of irrelevant objects (like leaf, branch), illumination differences due to weather, season and shadows. Therefore, features must be invariant to all above changes. Also, the model has to work with candidate animal patches that are of variable sizes and ratios since they are obtained through ensemble graph cuts. To handle above challenges, it present the following scheme f+ or animal– background verification model. 3Our scheme has three steps: (1) pre-processing, (2) fine-tuned DCNN features, and (3) classification through learning algorithms.

Existing literature has shown that DCNN is very efficient descriptors for object recognition, classification and retrieval, etc. There are multiple convolution layers and at least one fully connected layer in a DCNN. For translation invariant features, DCNN has pooling layer. Using the VGG-F pre trained model. The pretrained model has been learned on huge auxiliary ILSVRC 2012 dataset. The pretrained model has an image size for input of 224 × 224 hence, it resize the images to 224 × 224, without considering its actual size and ratio. The image resize incurs image distortions which can be neglected due to the fact that all the images go through the same distortions, and the effect of resizing is negligible. The DCNN provides a feature vector of 1000 dimensions.
This use DCNN features as they are self-learned features that enhance the performance of the system, and these features contain information that describes components of an image like edge, shape. The architecture of the VGG-F model is described in

detail .The parameters of each layer are given as convolution layer; number of filters with their size; stride value; spatial padding and down-sampling factor of max-pooling. Stride tells the allocation of spatial dimensions over the input while padding tells the size of the padding along the borders of the input in a convolution layer. Also, the pooling layer along with a convolution layer helps in reducing the size of the representations. Moreover, pooling aids in overcoming the problem of over fitting. Similarly for fully connected layers, the dimensionality of each layer along with the method used for regularisation is given, and in last layer, soft-max classifier is used that evaluates the deviation of output to the target.

IV. SYSTEM IMPLEMENTATION

Wild animal detection is reliable and robust method for animal detection in highly cluttered images using DCNN. The cluttered images are obtained using camera-trap networks. The images in camera-trap image sequences also provide the candidate animal region proposals done by multilevel graph cut. It is introduced by a verification step in which the proposed region is classified into animal or background classes, Thus, determining whether the proposed region is truly animal or not. We applied DCNN features to machine learning algorithm to achieve better performance. The experimental results shows that proposed system is efficient and robust wild animal detection system for both daytime and night time.

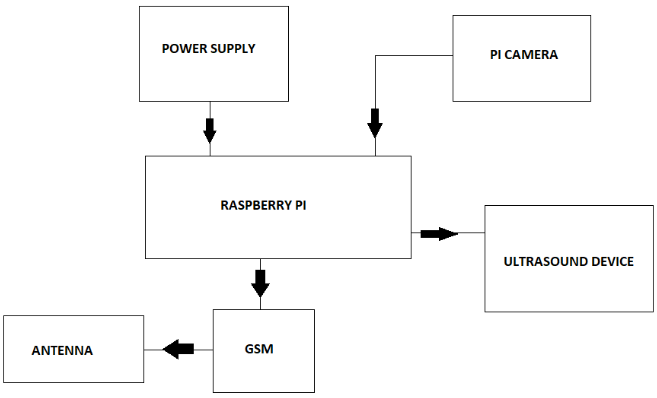

Figure 2 Block Diagram

Deep Convolutional Neural Network (CNN) is a special type of Neural Networks, which has shown exemplary performance on several competitions related to Computer Vision and Image Processing. Some of the exciting application areas of CNN include Image Classification and Segmentation, Object Detection, Video Processing, Natural Language Processing, and Speech Recognition. The powerful learning ability of deep CNN is primarily due to the use of multiple feature extraction stages that can automatically learn representations from the data. The availability of a large amount of data and improvement in the hardware technology has accelerated the research in CNNs, and recently interesting deep CNN architectures have been reported. Several inspiring ideas to bring advancements in CNNs have been explored, such as the use of different activation and loss functions, parameter optimization, regularization, and architectural innovations. However, the significant improvement in the representational capacity of the deep CNN is achieved through architectural innovations. Notably, the ideas of exploiting spatial and channel information, depth and width of architecture, and multi-path information processing have gained substantial attention. Similarly, the idea of using a block of layers as a structural unit is also gaining popularity. This survey thus focuses on the intrinsic taxonomy present in the recently reported deep CNN architectures and, consequently, classifies the recent innovations in CNN architectures into seven different categories. These seven categories are based on spatial exploitation, depth, multi-path, width, feature-map exploitation, channel boosting, and attention. Additionally, the elementary understanding of CNN components, current challenges, and applications of CNN are also provided.

4.1 CNN COMPONENTS

Nowadays, CNN is considered as one of the most widely used ML technique, especially in vision-related applications. CNN can learn representations from the grid-like data, and recently it has shown substantial performance improvement in various ML applications. Since CNN possesses both good feature generation and discrimination ability, therefore in a typical ML system, CNN capabilities are exploited for feature generation and classification. A typical CNN architecture generally comprises alternate layers of convolution and pooling followed by one or more fully connected layers at the end. In some cases, a fully connected layer is replaced with a global average pooling layer. In addition to different mapping functions, different regulatory units such as batch normalization and dropout are also incorporated to optimize CNN performance. The arrangement of CNN components plays a fundamental role in designing new architectures and thus achieving enhanced performance. This section briefly discusses the role of these components in a CNN architecture.

4.2 CONVOLUTION LAYER

The convolutional layer is composed of a set of convolutional kernels where each neuron acts as a kernel. However, if the kernel is symmetric, the convolution operation becomes a correlation operation. Convolutional kernel works by dividing the image into small slices, commonly known as receptive fields. The division of an image into small blocks helps in extracting feature motifs. Kernel convolves with the images using a specific set of weights by multiplying its elements with the corresponding elements of the receptive field. Due to weight sharing ability of convolutional operation, different sets of features within an image can be extracted by sliding kernel with the same set of weights on the image and thus makes CNN parameter efficient as compared to the fully connected networks. Convolution operation may further be categorized into different types based on the type and size of filters, type of padding, and the direction of convolution.

4.3 POOLING LAYER

Feature motifs, which result as an output of convolution operation, can occur at different locations in the image. Once features are extracted, its exact location becomes less important as long as its approximate position relative to others is preserved. Pooling or down-sampling is an interesting local operation. It sums up similar information in the neighborhood of the receptive field and outputs the dominant response within this local region. The use of pooling operation helps to extract a combination of features, which are invariant to translational shifts and small distortions. Reduction in the size but also helps in increasing the generalization by reducing overfitting. Different types of pooling formulations such as max, average, L2, overlapping, spatial pyramid pooling, etc. are used in CNN.

4.4 TRAINING OF CNN

An SGD optimizer was used to optimize CNN training. In the SGD optimizer, optimization is performed using a step policy that multiplies the gamma value for each fixed iteration so that the training accuracy and loss converge quickly. Training, a function of SGD, was performed in mini-batch size units. The number of iterations is calculated as "number of training data / mini-batch size," defined as 1 epoch. In this experiment, learning rate is 0.0005, momentum is 0.9, and gamma is 0.1, whereas the minibatch size in ResNet-50 is 16, 6 in ResNet-101, 3 in ResNet-152, and 20 in visual geometry group (VGG)-16, with a maximum epoch of 10. Because fine-tuning was performed using the existing pre-trained weights, we used small learning-rate values. After seven epochs, the learning rate is reduced. One epoch indicates that training is performed as many times as the total number of iterations.

Therefore, the total number of trainings is equal to the number of iterations × the number of epochs. To calculate the training loss, the softmax function was used to calculate multinomial logistic loss. When training is performed, the accuracy converges to 100 and the loss converges to 0. This shows that the training of the CNN model used in this study was successful. We made the self-collected DMFW-DB1 and trained ResNet model available to other researchers through for fair comparisons.

4.5 COMPARISON OF FINGER-WRINKLE RECOGNITION PERFORMANCE ACCORDING TO COLOR SPACE

For the first experiment, we compared the recognition performance of input images of gray and various color spaces. In the following experiment, we compared the recognition performance of Retinex filtering and the original image. but it also increases the distinctiveness of finger-wrinkle texture. Moreover, the deep ResNet used in this study can obtain sufficiently robust features to address the illumination variation. In the next experiment, we measured the recognition accuracy according to the processing method of the background region.

V. RESULT AND ANALYSIS

DCNN Algorithm is processed and the animal image is detected and classified by comparing the input image with existing dataset. To check the working of the exact output the code is done and the sample input is given as chair.

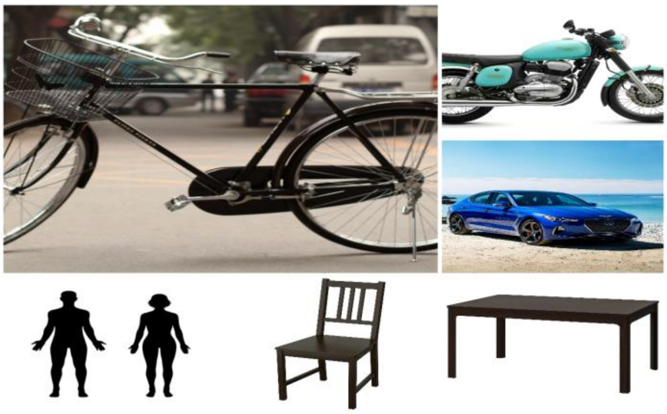

Figure 3 Input-1 for animal identification

Figure 4 Detection Output

Figure 5 Input-2 for animal identification

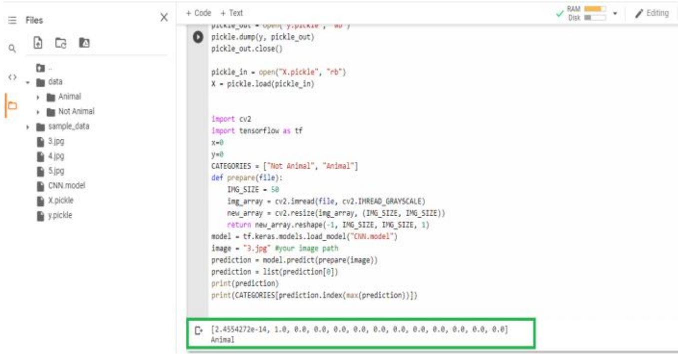

Figure 6 Detection Output

VI. CONCLUSION

Deep Convolution Neural Network (DCNN) algorithm is used to detect wild animals. DCNN algorithm classifies animals efficiently with a good number of accuracy and also the image of the detected animal is displayed for a better result so that it can be used for other purposes such as detecting wild animals entering into human habitat and to prevent wildlife poaching then gives an alarm sounds and send message to farmers. Also, the future scope of this system can be extended to various purposes like roadside safety of traveller from wild animals in forest bound roads. In future along with identification of wild animals, control measures like sedating based on animal type can be implemented with more real time data and advanced sensors and systems which would make this technique completely automated without any manual interference.

VII. REFERENCES

[1]. Fang, Y., Du, S., Abdoola, R., Djouani, K., & Richards, C. (2016). Motion based animal detection in aerial videos.Procedia Computer Science, 92, 13-17.

[2]. Jasko, G., Giosan, I., &Nedevschi, S. (2017, September). Animal detection from traffic scenarios based on monocular color vision. In Intelligent Computer Communication and Processing (ICCP), 2017 13th IEEE International Conference on (pp. 363-368). IEEE.

[3]. Nguyen, H., Maclagan, S. J., Nguyen, T. D., Nguyen, T., Flemons, P., Andrews, K., ... & Phung, D. (2017, October). Animal recognition and identification with deep convolution neural networks for automated Animal monitoring. In Data Science and Advanced Analytics (DSAA), 2017 IEEE International Conference on (pp. 40-49). IEEE.

[4]. Kumar. S., & Singh, S. K. (2016). Monitoring of pet animal in smart cities using animal biometrics. Future Generation Computer Systems.

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

[6]. Matuska, S., Hudec, R., Kamencay, P., Benco, M., &Zachariasova, M. (2014). Classification of wild animals based on SVM and local descriptors. AASRI Procedia, 9, 25-30.

[7]. Villa, A. G., Salazar, A., & Vargas, F. (2017). Towards automatic wild animal monitoring: Identification of animal species in camera-trap images using very deep convolutional neural networks. Ecological informatics, 41, 24-32.

[8]. Xue, C., Wang, P., Zhao, J., Xu, A., & Guan, F. (2017). Development and validation of a universal primer pair for the simultaneous detection of eight animal species. Food chemistry, 221, 790-796.

[9]. Xue, W., Jiang, T., & Shi, J. (2017, September). Animal intrusion detection based on convolutional neural network. InCommunications and Information Technologies (ISCIT), 2017 17th International Symposium on (pp. 1-5). IEEE.

[10]. Zhang, T., Wiliem, A., Hemsony, G., & Lovell, B. C. (2015, April). Detecting kangaroos in the wild: the first step towards automated animal surveillance. In ICASSP (pp. 1961-1965).

[11]. Zhu, C., Li, T. H., & Li, G. (2017, October). Towards automatic wild animal detection in low quality camera-trap images using two-channeled perceiving residual pyramid networks. In Computer Vision Workshop (ICCVW), 2017 IEEE International Conference on (pp. 2860-2864). IEEE.

Cite this article as :

Meenatchi K, Thibishini V, Vaisnavi K, Mrs. R. Ahila, "Wild Animal Detection System Using Deep Convolutional Neural Networks ", International Journal of Scientific Research in Science and Technology (IJSRST), Online ISSN : 2395-602X, Print ISSN : 2395-6011, Volume 8 Issue 3, pp. 989-996, May-June 2021.

Journal URL: https://ijsrst.com/IJSRST2183215