

# Integrated Animal Recognition and Detection Using Deep Convolutional Neural Network

G.Supriya<sup>1</sup>, Mr. A.S.Loganathan M.Tech.,<sup>2</sup>, Mr. K.Balaji M.E.,<sup>3</sup>P.G. Student, Department of ECE, GRT Institute of Engineering and Technology, Tiruvallur, Tamil Nadu, India<sup>1</sup>Assistant Professor, Department of ECE, GRT Institute of Engineering and Technology, Tiruvallur, Tamil Nadu, India<sup>2,3</sup>

**ABSTRACT:** In this paper, the Deep Convolutional Neural Network (DCNN) for the classification of the input animal images is proposed. Wildlife monitoring and analysis is 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 neighbor, and ensemble tree. The overall performances were obtained using different number of training images and test images. The experimental results show that the proposed method has a positive effect on overall animal recognition performance and outperforms other examined methods.

**KEYWORDS:** Age detection , DCNN , State of art

## I. INTRODUCTION

Animals can be recognized using their footprints. Some features contained within an animal footprint can be utilized to aid in the recognition of an animal. Amongst these features, the most general and most utilized by humans to manually identify the animal is the number and size of blobs in the footprint. Using image processing techniques an algorithm was created to segment and extract the best possible representation of the footprint which varied across colour. Connected Components was then utilized to count the number of blobs contained within the footprint and measure the size of every blob. The proposed system is used to predict the animal's age based on the footprints.

Researches regarding animals in image processing have been an important field to numerous applications. Many algorithms and methods have been developed by human being in order to have a better understanding on animal behaviour. Besides, these applications also can act as a warning system to human being from intrusion of dangerous wild animal for early precaution measures. These applications can be narrowed down to three main branches, namely detection, tracking and identification of animal. The first branch, which is the animal detection, has been applied in various fields of real life application. For example, a detection algorithm has been developed for light detection and ranging (LIDAR) data to enable fisherman to find the right location of fishes in deep sea [1]. Researchers in [2] used airborne remote sensing images to detect moving wild animals. Work in [3] on the other hand, used micro-Doppler signals to maintain the human security by detecting potential dangerous animal intrusion to residential area. The second branch, which is the animal tracking, is the main topic in monitoring animal locomotive behaviour and its interaction with the environment. With the technology of sensor, radio-frequency identification (RFID), and global positioning system (GPS), one of the applications is the development of new zoological systems for animal trace ability, identification, and anti-theft for the management and security of animal in zoo . By tracking the animal movements, it helps human to have a better understanding on living creatures on earth, especially on how the animal interacts with its

# International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: [www.ijirset.com](http://www.ijirset.com)

Vol. 8, Issue 4, April 2019

environment. The third branch, which is the animal identification, is used to identify the detected animal. It has been used in health monitoring system for domestic herds.

## II. RELATED WORK

### **Small-Footprint Highway Deep Neural Networks for Speech Recognition Li-Liang Lu ; Steve Renals**

In this paper, we study the application of the recently proposed highway deep neural network (HDNN) for training small-footprint acoustic models. HDNNs are a depth-gated feed forward neural network, which include two types of gate functions to facilitate the information flow through different layers. Our study demonstrates that HDNNs are more compact than regular DNNs for acoustic modeling, i.e., they can achieve comparable recognition accuracy with many fewer model parameters.

### **A Privacy-Aware User Tracking System Using Footprint Data Measured across Multiple Pressure Sensor Sheets - Ryoya Wada, Shigeaki Tagashira, Masaki Ogino, Shigemi Ishida, Akira Fukuda**

In this paper, we propose a privacy-aware user tracking system which traces users by their footprints. Tracking by footprints could reduce discomfort for privacy-sensitive users compared with the use of face recognition by cameras in existing tracking systems, since a footprint is weakly identifiable personal information, i.e., it cannot directly identify a specific individual. In this context, the proposed system is said to be privacy-aware. In the proposed tracking system, the footprint data are collected from spatially disjoint pressure sensor sheets put on significant locations in a target area.

### **Infant Footprint Recognition - Eryun Liu**

In this paper, we investigate the feasibility of infant recognition by their footprint using a 500 ppi commodity friction ridge sensor. We collected an infant footprint dataset in three sessions, consisting of 60 subjects, with age range from 1 to 9 months. We proposed a new minutia descriptor based on deep convolutional neural network for measuring minutiae similarity. The descriptor is compact and highly discriminative. We conducted verification experiments for both single enrolled template and fusion of multiple enrolled templates, and show the impact of age and time gap on matching performance. Comparison experiments with state of the art algorithm show the advantage of the proposed minutia descriptor.

## III. PROPOSED ALGORITHM

Design a framework for animal recognition in the wild, aiming at a fully automatic wildlife spotting system. Our work is motivated by the state-of-the-art power of recent deep CNN models for image classification, in particular the recent evidence that automated recognition can surpass human at certain object recognition tasks in the ImageNet competition. A multiclass classifier to label each input image with animal presence by a specified species. The core of each task is essentially a deep CNN-based classifier, trained from prepared datasets manually labeled by volunteers. Several selected deep CNN architectures are employed to the framework for comparisons.

Our proposed image classification framework and its application to the Wildlife Spotter datasets. First we describe the datasets. Then we introduce a CNN based framework for wildlife identification. Finally we characterize selected CNN architectures employed in our experiments and implementations.

For each task we train CNN models in two scenarios: imbalanced and balanced datasets. We compute classification accuracy for both cases. In case of dataset imbalance, Fmeasure is employed in addition to accuracy, to test the robustness of the proposed system. In machine learning, a convolutional neural network (CNN, or ConvNet) is a class of deep, feed-forward artificial neural networks, most commonly applied to analyzing visual imagery. CNNs use a variation of multilayer perceptrons designed to require minimal preprocessing. They are also known as shift invariant or space invariant artificial neural networks (SIANN), based on their shared-weights architecture and translation invariance characteristics. Convolutional networks were inspired by biological processes in that the connectivity pattern between neurons resembles the organization of the animal visual cortex. The feature extraction stage consists of

# International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: [www.ijirset.com](http://www.ijirset.com)

Vol. 8, Issue 4, April 2019

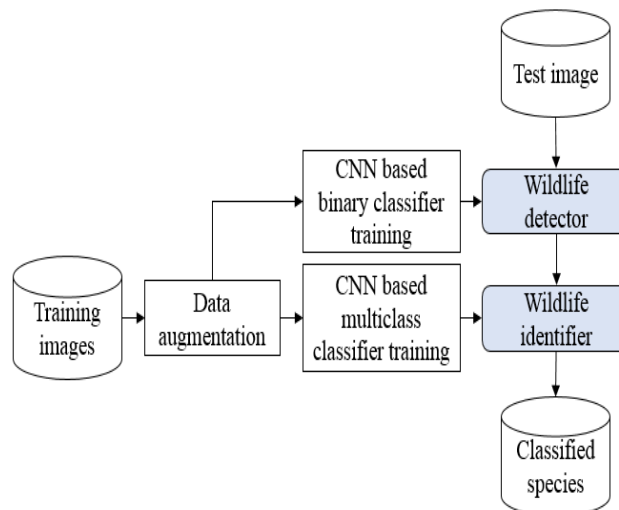
alternating trainable fully convolutional (conv) layers which learn high-level features while the classifier part, which is the fully connected layer, deals with different combinations of those features in order to make the final decisions.

## Image Processing

Digital image processing is the use of computer algorithms to perform image processing on digital images. As a subcategory or field of digital signal processing, digital image processing has many advantages over analog image processing. It allows a much wider range of algorithms to be applied to the input data and can avoid problems such as the build-up of noise and signal distortion during processing. Since images are defined over two dimensions (perhaps more) digital image processing may be modeled in the form of multidimensional systems.

## Gray Image

A grayscale or greyscale image is one in which the value of each pixel is a single sample representing only an amount of light, that is, it carries only intensity information. Images of this sort, also known as black-and-white or gray monochrome, are composed exclusively of shades of gray. The contrast ranges from black at the weakest intensity to white at the strongest.



## Algorithm description

The main idea of the median filter is to run through the signal entry by entry, replacing each entry with the median of neighboring entries. The pattern of neighbors is called the "window", which slides, entry by entry, over the entire signal. For 1D signals, the most obvious window is just the first few preceding and following entries, whereas for 2D (or higher-dimensional) signals such as images, more complex window patterns are possible (such as "box" or "cross" patterns). Note that if the window has an odd number of entries, then the median is simple to define: it is just the middle value after all the entries in the window are sorted numerically.

## Advantages

- This independence from prior knowledge and human effort in feature design is a major advantage.
- More efficient and
- Efficiency in hardware implementation performance metric.

# International Journal of Innovative Research in Science, Engineering and Technology

*(A High Impact Factor, Monthly, Peer Reviewed Journal)*

Visit: [www.ijirset.com](http://www.ijirset.com)

Vol. 8, Issue 4, April 2019

## IV. RESULTS AND DISCUSSIONS

The experimental results have shown that high accuracy can be achieved in detecting images that contain animals with more than 96% accuracy. For citizen science-based projects and systems, human annotators' time are valuable and this framework has the promise to radically improve the efficiency in these systems. For example, for the current Wildlife Spotter system up to the time we collected this dataset, more than 32% of images presented to human annotators does not contain any animal to be annotated. The animal identification results have shown a good performance in identifying the most common three and six species of animals. While this performance may not yet be sufficient to build a fully automatic recognition, it still adds an enormous value in improving the system by automatically providing initial animal labels for human annotators. We anticipate that with more data collected over time and with a rapid growing capacity of deep learning techniques in computer vision, this performance could be improved significantly in near future

original image



The above figure shows the foot print of animal. And the above image is further converted into angray scale image shown in 2<sup>nd</sup> figure.

Grayscale image



# International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: [www.ijirset.com](http://www.ijirset.com)

Vol. 8, Issue 4, April 2019

THRESHOLDED IMAGE

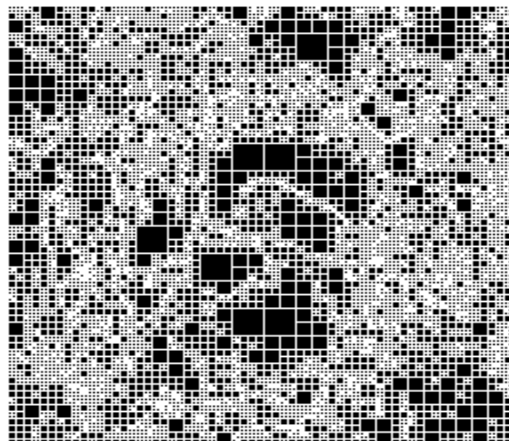


The above figure shows the Thresholding of the gray scale image

The input to a thresholding operation is typically a grayscale or color image. In the simplest implementation, the output is a binary image representing the segmentation. Black pixels correspond to background and white pixels correspond to foreground (or vice versa). In simple implementations, the segmentation is determined by a single parameter known as the intensity threshold. In a single pass, each pixel in the image is compared with this threshold. If the pixel's intensity is higher than the threshold, the pixel is set to, say, white in the output. If it is less than the threshold, it is set to black.

In more sophisticated implementations, multiple thresholds can be specified, so that a band of intensity values can be set to white while everything else is set to black. For color or multi-spectral images, it may be possible to set different thresholds for each color channel, and so select just those pixels within a specified cuboid in RGB space. Another common variant is to set to black all those pixels corresponding to background, but leave foreground pixels at their original color/intensity (as opposed to forcing them to white), so that that information is not lost.

QUAD TREE DECOMPOSITION IMAGE



The above image shows the Convolution Quad Tree Decomposition image.

## International Journal of Innovative Research in Science, Engineering and Technology

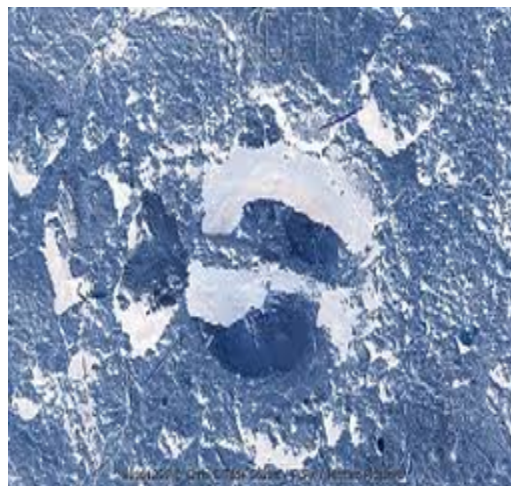
(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: [www.ijirset.com](http://www.ijirset.com)

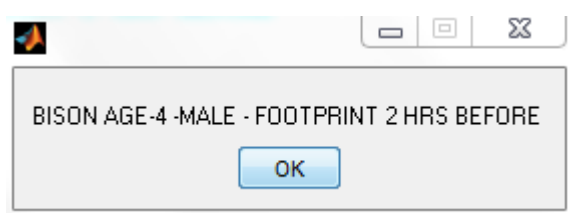
Vol. 8, Issue 4, April 2019

There are several advantages of using a quadtree to hierarchically represent the image data for convolutional neural networks. First, the quadtree can be efficiently computed once for each image, and its structure then remains fixed throughout the forward and backward passes and requires no further bookkeeping. Its structure defines the locations of the non-zero pixels hierarchically and the nodes that are non-empty in each level. Convolutions are performed on the pixels that correspond to the nodes in the bottommost level of the quadtree, resulting in feature maps that fit in the same level. By simply restricting computations to the sparse quadtree nodes, we can ensure that feature maps do not dilate in the deeper layers of the network even when repeated convolution layers are stacked, hence retaining the sparse nature of the input. Second, since the quadtree structure is by definition hierarchical, downsampling and upsampling features can be performed easily and efficiently. Pooling downsamples the feature map such that it can be stored in the previous level of the quadtree, and is carried out by assigning the maximum or average of the children nodes at the current level into their parent node in the previous coarser level. Upsampling can be performed similarly by traversing the quadtree in the opposite direction

color segmented Image



The above figure shows the Color Segmented Image. The most widely used color space is the color space, where a color point in the space is characterized by three color components of the corresponding pixel which are red (), green (), and blue (). However since there exist a lot of color spaces, it is useful to classify them into fewer categories with respect to their definitions and properties. Image segmentation is an essential but critical component in low-level vision, image analysis, pattern recognition, and now in robotic systems. Besides, it is one of the most difficult and challenging tasks in image processing, and determines the quality of the final results of the image analysis. Intuitively, image segmentation is the process of dividing an image into different regions such that each region is homogeneous while not the union of any two adjacent regions. An additional requirement would be that these regions had a correspondence to real homogeneous regions belonging to objects in the scene.



The above image is final classified image.

# International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: [www.ijirset.com](http://www.ijirset.com)

Vol. 8, Issue 4, April 2019

## V. CONCLUSION AND FUTURE WORK

A system for automatically classifying animal footprints by first classifying them by means of the connected components algorithm, and then using morphological features to further classify them into a correct species of animal was created. 3 Distance formulas and a combination of all three formulas were used to make classifications. The algorithm boasted a 97% final accuracy rate and a perfect 100% accuracy rate for initial classification. The age of the animal is also predicted.

## REFERENCES

1. P. M. Vitousek, H. A. Mooney, J. Lubchenco, and J. M. Melillo, "Humandomination of Earth's ecosystems," *Science*, vol. 277, no. 5325, pp. 494–499, 1997.
2. G. C. White and R. A. Garrott, *Analysis of wildlife radio-tracking data*. Elsevier, 2012.
3. R. Szewczyk, A. Mainwaring, J. Polastre, J. Anderson, and D. Culler, "An analysis of a large scale habitat monitoring application," in *Proceedings of the 2nd International Conference on Embedded Networked Sensor Systems*, 2004, pp. 214–226.
4. B. J. Godley, J. Blumenthal, A. Broderick, M. Coyne, M. Godfrey, L. Hawkes, and M. Witt, "Satellite tracking of sea turtles: Where have we been and where do we go next?" *Endangered Species Research*, vol. 4, no. 1-2, pp. 3–22, 2008.
5. A. Hulbert and J. French, "The accuracy of GPS for wildlife telemetry and habitat mapping," *Journal of Applied Ecology*, vol. 38, no. 4, pp. 869–878, 2001.
6. R. Kays, S. Tilak, B. Kranstauber, P. A. Jansen, C. Carbone, M. J. Rowcliffe, T. Fountain, J. Eggert, and Z. He, "Monitoring wild animal communities with arrays of motion sensitive camera traps," arXiv:1009.5718, 2010.
7. F. O'Connell, J. D. Nichols, and K. U. Karanth, *Camera traps in animal ecology: Methods and Analyses*. Springer Science & Business Media, 2010.
8. Swanson, M. Kosmala, C. Lintott, R. Simpson, A. Smith, and C. Packer, "Snapshot Serengeti, high-frequency annotated camera trap images of 40 mammalian species in an African savanna," *Scientific Data*, vol. 2, p. 150026, 2015.
9. S. Thorpe, D. Fize, and C. Marlot, "Speed of processing in the human visual system," *Nature*, vol. 381, no. 6582, p. 520, 1996.
10. X. Yu, J. Wang, R. Kays, P. A. Jansen, T. Wang, and T. Huang, "Automated identification of animal species in camera trap images," *EURASIP Journal on Image and Video Processing*, vol. 2013, no. 1, pp. 1–10, 2013.
11. G. Chen, T. X. Han, Z. He, R. Kays, and T. Forrester, "Deep convolutional neural network based species recognition for wild animal monitoring," in *Proceedings of the IEEE International Conference on Image Processing (ICIP)*, 2014, pp. 858–862.
12. Gómez, A. Salazar, and F. Vargas, "Towards automatic wild animal monitoring: Identification of animal species in camera-trap images using very deep convolutional neural networks," arXiv:1603.06169, 2016.
13. J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, "ImageNet: A large-scale hierarchical image database," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2009, pp. 248–255.
14. O. Russakovsky, J. Deng, H. Su et al., "ImageNet large scale visual recognition challenge," *International Journal of Computer Vision*, vol. 115, no. 3, pp. 211–252, 2015.
15. N. Pinto, D. D. Cox, and J. J. DiCarlo, "Why is real-world visual object recognition hard?" *PLOS Computational Biology*, vol. 4, no. 1, p. e27, 2008.
16. M. Bishop, "Pattern recognition," *Machine Learning*, vol. 128, pp. 1–58, 2006.
17. Y. LeCun, B. Boser, J. S. Denker, D. Henderson, R. E. Howard, W. Hubbard, and L. D. Jackel, "Backpropagation applied to handwritten zip code recognition," *Neural Computation*, vol. 1, no. 4, pp. 541–551, 1989.
18. K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," arXiv:1409.1556, 2014.