Animal Detection Using Histogram Oriented Gradient

Mukesh B.Rangdal
ME VPCOE Baramati
Pune India.
mukeshrangdal@gmail.com

Dinesh B. Hanchate
Asst. Prof. VPCOE Baramati
Pune India.
dineshbhancahte@gmail.com

Abstract— Animal detection are useful for many real life applications. Animal detection methods are useful on research related identification of targeted animal and also prevented wild life animal intrusion in residential area. The animal detection becoming important issue because of the accidents occurring in the residential area. For preventing this issue, the AHDS system is proposed. The system consists animal detection using HOOG - [Haar of Oriented Gradients]. The HOOG is used for extracting the shape and texture features separately and combined those features using joint Learning approach for animal detection. After combining the shape and texture feature, two detection algorithms have been used for identifying the animals.

Keywords- *Classifiers, HOG processing, gradients, Haar of Oriented Gradients.*

*****

INTRODUCTION

Due to the rapid concretization of the lands and urbanization, the shelter places of the animals, birds etc. are getting diminished and these creatures are finding it very difficult to survive and some species are becoming extinct. To survive life of animal intrusion in residential area from residential people to get inform to prevent an accidental hurt to wild life and save both civilian and wild life. This system prevents the intruders from entering into the restricted area. Moreover the detection must be done on animals. This is so because in many cases the animals hurt themselves by the residential people. To prevent this, one detects animal if in the case of to prevent the animals from harm. There are so many human detectors which are capable of detecting only humans. But those detectors cannot be used for animal detection. Since the humans and animals shapes and texture varies the human detectors cannot be simply applicable to animals. This urged to find the detector for the animal. The process is initiated once the object captured by the camera in the surveillance area.

I. LITERATURE SURVEY

A. Like features

In[11] Viola and Jones adapted a technique in rapid object detection which is known as the Haar like features. These techniques are adopted in the digital image feature for object recognition. The haar wavelets are the one which is used in the real time human face detector and this method has some similarity with that. The working with image intensities is quite costly. Viola and Jones adapted the techniques from haar wavelet to form haar like features.

The process of this get adjacent rectangular region at a location and calculating the sum of the pixel intensity values in that region and then the difference is found. For the detection process the object of ridicule is moved over the source image and for each subsection the haar like feature value is found. The difference found is used to learn about the image. Since this is a weak classifier at large, It need a large set of training samples. The advantage of this is the calculation speed and at constant time the image of any size can be calculated. This method cannot be used for animal detection since the shape and texture varies.[11].

B. Wavelet coefficients

The wavelet coefficients are similar to the form of a wave. These initiate from the zero and then reach the peak and once again reach zero. For the purpose of visualization this can be found in the heart monitor. The wavelet has some identical properties which makes it useful in signal processing. The wavelets are combined to extract the designated signal using the source signal. The concept of the resonance is the core of many practical applications of the wavelet concepts. These techniques can be used for human detection because of the performance and not for animals because the shape and texture plays an important role.

C. Scale in-variant feature transform

In Lowe [7], David Lowe had identified and described the local features in an image. The description from the training image is used to identify the object for locating the test image from the given samples. The image must be detected even when there is change in the image scale, noise and illumination. These points lie as object edges. The relative positions between the features in the source scene will not change from one image to another. Features present in the stretchy objects will not work if the changes are present in the internal geometry. SIFT is effective in detection since this uses a large number of features from the images and there by reduces the feature matching errors. But this method cannot be used because, accuracy and shape and texture are important for animal detection. The above methods are used only for human posture detection, not for animal detection. In order to overcome this problem, the proposed AHDS system incorporates animal detection.

Feature extraction Methods:

Method	Notes
Principal Component Analysis (PCA)	Eigenvector-based, linear map
Kernel PCA	Eigenvector-based , non-linear map, uses kernel methods
Weighted PCA	PCA using weighted coefficients
Linear Discriminant Analysis (LDA)	Eigenvector-based, supervised linear map
Semi-supervised Discriminant (SDA)	Analysis Semi-supervised adaptation of LDA
Independent Component Analysis (ICA)	Linear map, separates non-Gaussian distributed features
Neural Network based methods	Diverse neural networks using PCA, etc.
Multidimensional Scaling (MDS)	Nonlinear map, sample size limited, noise Sensitive
Self-organizing map (SOM)	Nonlinear, based on a grid of neurons in the feature space
Active Shape Models (ASM)	Statistical method, searches boundaries
Active Appearance Models (AAM)	Evolution of ASM, uses shape and texture
Gavor wavelet transforms	Biologically motivated, linear filter
Discrete Cosine Transform (DCT)	Linear function, Fourier-related transform, usually used 2D-DCT
MMSD, SMSD	Methods using maximum scatter difference criterion.

Table 1: Feature extraction algorithms

II. PROPOSED METHOD

A. HOG Algorithm[16]

- The technique is based on the evaluation of histograms calculated on the basis of the orientation of the gradients of the input image. The concept behind is the edges and the shapes of objects can be well characterized by the intensity of individual local gradients. Computation of each histogram is obtained by dividing the image into a grid of cells and for each of them is computed a histogram of gradients (sobel 3X3 operator). Then the cells are grouped into regions called blocks. In addition, to release the response[16]
- From information of the individual blocks you get a descriptor (trained dataset in vector format) that will be used for detection. The above operations are performed on windows of finite size that analyze the image at multiple scales. For each window is then obtained a descriptor that is then given to a linear SVM classifier that provides predictions about the presence or absence of a Animal.[16]
- Being the construction procedures of individual descriptors independent of each other, you can run them simultaneously with a multi-threading system. In this way you get a significant performance improvement. The robustness of the algorithm used (SVM algorithm) and accurate image scanning at multiple scales, often produce a dense group of detection windows for each object, then a merger is required (mean-shift) that brings up to a final identification window. by the terms of brightness, can be useful normalize individual blocks. [16]

B. Peculiarity of multi scale HOG: Dalal and Triggs[3] proposed the Histogram of Oriented Gradients which is successfully used in the human detection. The difference was made in training the sample images[14]. The pixel value for the images was 64X128. The windows were divided by 8X8 cell. The multi scale HOG refers to the 36-dimensional feature of each block. Block contains four cells. Each cell has 9-dimensional feature vector. The feature of HOG is the shape feature of the gradient image. This method uses the Adaboost mechanism for training the classifiers as strong classifiers. The gradient image of the training sample is found and then the blocks are selected for finding the characteristics of the gradient image [15].

III. EXPRIMENTAL RESULT

The organization of the project, as seen in figure 1 clearly separates different parts of the program. The Core contains the main processing like image acquisition, pre-processing, feature extraction. The UI is the Graphical User Interface which presents the Image to the user and provides some features that can modify the Core behaviour. The HOG contains some tools for the Histogram of Oriented Gradients computation. The Utils part are functionalities which are useful for several parts of the system.

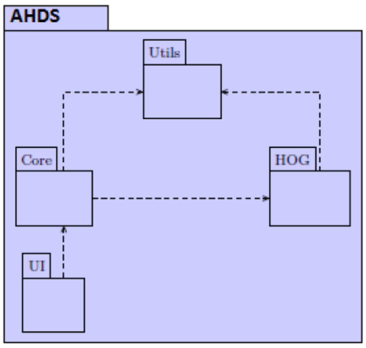


Figure 1: General presentation

1) Animal Training Set Generation

In order to train both of these detectors, we needed a large training data set of cropped cat face images .Luckily, the authors of [1] created a database of 10,000 cat images where

the faces in each image were hand labeled. However, these images consisted of the entire cat in various settings with varied image sizes. On top of that, there was no guarantee that the cat faces where horizontal since these images were extracted from the Internet. Therefore, we needed to crop each of these images so they were centered on a horizontal
face at our chosen resolution for training. Because hand cropping 10,000 images would take too much time, we wrote a java code to automatically crop the images. This is
done as follows:

a) Using the labeled eye and ear points, we calculate a rotation angle so that these features are horizontal.

b) Locate the center of the face by calculating the centroid of the point set (excluding ear tips and double counting the nose point).

c) Estimate the crop region by calculating the height (distance between the mouth and the line connecting ear *tips) and width (distance between ear tips) and create a box* that is 1.1 times the height and width of the cat face, centered at the centroid calculated..

d) Extract the boxed region, convert it to gray scale and then resize it to our 32 x 32 training set image size It chose an image size

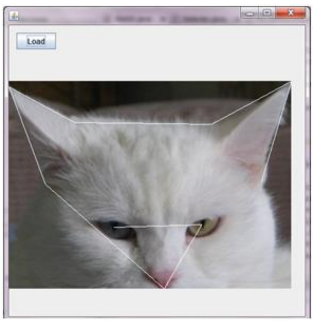


Figure 2: Example of label and croped image

B. SVM Classifier

A Support Vector Machine (SVM) is a machine learning technique used for classification or regression2. It was introduced by Vladimir N. Vapnik and C. Cortes in 1995 [12]. In other words it is a technique that allows to organize data according classes (or category) in order to determine the class of another sample.
Definition In a n-dimensional space, an hyperplane is a subset of dimension n 1 which separates the space into two half spaces. It can be described as a set of points x
such as.

$$w.x - b = 0 \qquad (1)$$

where w is the normal vector to the hyperplane. In the following, the algorithm uses a set D of n elements of class y as a training set.

$$d = \{(x_i, y_i | x_i \varepsilon R^P, y_i \varepsilon \{-1,1\}\}_{i=1}^{n} \qquad (2)$$

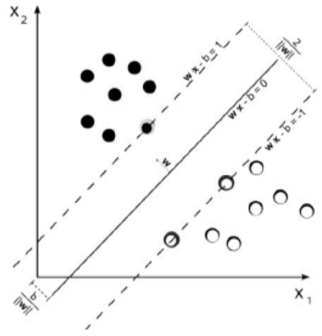


Figure 4: Description of Hyperplanes and Margins for Support Vector Machines. Plain line represent the maximum-margin hyperplane. Dashed lines represent the margins. Bold dots are the Support Vectors.

C. Histogram Oriented Gradients

Histogram of Oriented Gradients, commonly called HOG, is a computer vision technique commonly used to detect all kind of objects. Unlike background subtraction algorithms,
it has a region-based approach which makes it robust against geometric and photometric transformations. The motivation for using this technique is mainly its robustness. Indeed, the training is computed with SVM on a database with hundreds of images. The algorithm is designed to recognize specific kinds of objects, which makes it a good choice for Animal detection purposes. Compared to other algorithm, like ASM (Active Shape Model),

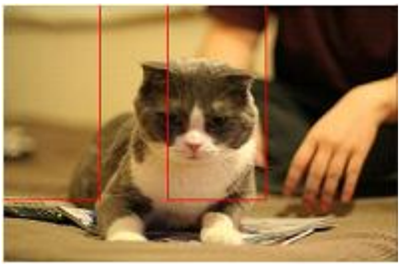


Figure :3 Example using HOG

which uses statistical models and deform the shape of an object to fit a reference image, HOG shows better result in both positive and negative matches (meaning that there are rate of true positive/ true negative is better). The algorithm was described by N. Dalal and B. Triggs in [3], It uses gradients magnitude and direction to describe all parts (or cells) of an image[9]. It can be combined with a SVM detector for human detection purposes. Figure 5 represents the different steps to compute the histogram of oriented gradients.

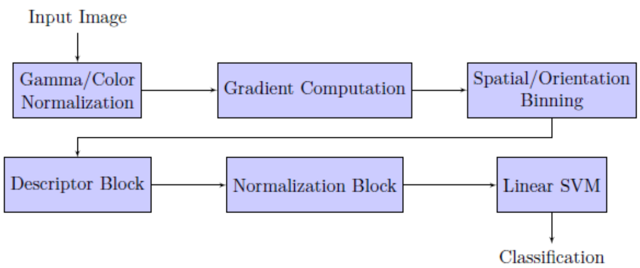


Figure 5: Fundamental HOG Algorithm

.

D. Gamma/Colour Normalization

The main purpose of the Gamma/Colour normalization block is to reduce the impact the light has on the given image (local shadowing and illumination variations). The used images should be in RGB or LAB colour spaces as the coloured images show better results than the gray-scale ones. The normalization itself does not improve performance significantly, probably because a similar job is done again later.[3]

E. Gradient computation

The importance of determining the gradients from one image is due to the fact they get information related to contours, silhouettes and sometimes texture. All these aspects are important when defining specific objects such as cars, animals or humans. The gradients are computed by applying some kind of derivative mask over the input image. However it was showed in [2] that the most effective technique for gradient computation is the application of a 1-D, centred, point discrete derivative mask to filter the colour or intensity of the image. The following masks are used:

$$[101] \text{ or } [101]^T$$

This gives us the gradient at each point which allows us to obtain the orientation and the magnitude; (using the $G_x$and G_y and the x and y projections of the gradient). The following formulas show the magnitude(3) and orientation(4):

$$G=\sqrt{G_x} + G_y \qquad (3)$$

$$\boldsymbol{\theta}=\arctan\frac{G_x}{G_y} \qquad (4)$$

F. Spatial/Orientation Binning

This is the part where the histogram of oriented gradients is built. Each pixel in the image holds a weighted vote for the orientation, based on the gradients computed in the previous step. The image is divided in small spatial regions called cells (see Figure 6), and for every cell the orientations of the gradients are accumulated in a 1D histogram with a predefined number of bins. Each pixel votes for one of the histograms bins according to its orientation, and increases it with its magnitude. These orientation bins

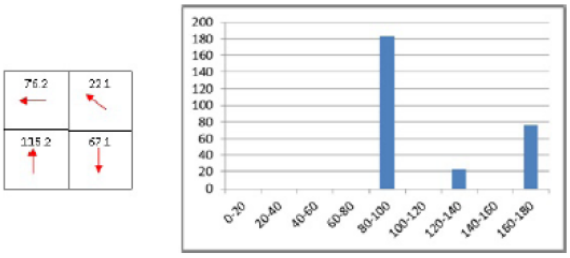


Figure 6: A Cell (with orientation and magnitude) and its Histogram

are evenly spaced over 0^o-180^o in case the gradient is " unsigned", or 0^o - 360^o in case the gradient is "signed". The "signed" gradient is of no use for human detection, but shows better results for other object detection. Usage shows that 9-class (9 bins) histograms are sufficient.

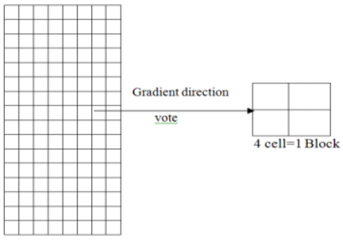


Fig.8: The training sample for the window specification[14]

G. Normalization and Descriptor Block

As mentioned earlier the normalization process helps in reducing the light variations or shadowing problems. In this step, cells are grouped into bigger units called blocks. Usually the blocks are overlapped and the cells are shared between the blocks and also normalized separately. the R-HOG(rectangular) block and the C-HOG (circular) block.
The R-HOG block is the default descriptor used in [3]. The R-HOGs are similar to SIFT descriptors (Scale Invariant Feature Transform) but they don't align to their dominant orientation. The R-HOGs are computed over m x m grids (m being the number of cells in each block) of n x n pixel cells and histogram bins. Figure 7 shows an R-HOG blocks composed of 3x3 cells of 6x6 pixels. However the best configuration is using a 2x2 cells of 8x8 pixels with 9 histogram bins. There are three methods in which the blocks can be normalized where v is the un normalised descriptor vector and e which is a normalisation constant that prevents the division by zero:

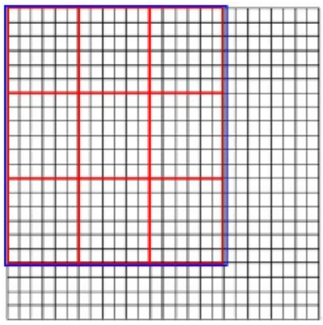


Figure 7: Representation of pixels (black), cells (red), and R-HOG block (blue).

$$\text{L2-norm:f}=\frac{v}{\sqrt{\|v\|_2^2+e^2}} \qquad (5)$$

$$\text{L1-norm:f}=\frac{v}{\|v\|_1+e^2} \qquad (6)$$

$$\text{L1-SQ:f}=\sqrt{\frac{v}{\|v\|_1+e^2}} \qquad (7)$$

IV Conclusion

This paper presented a Animal head detection system. good results could be achieved by decomposing texture and shape features firstly and fusing detection results secondly. The texture and shape detectors also greatly benefit from a set of new oriented gradient features. Although the focus is on the cat head detection problem in this paper, the approach can be extended to detect other categories of animals. In the future, the approach can be extended to multi-view cat head detection and more animal categories.

REFERENCES

[1] Weiwei, Z., S. Jian and T. Xiaoou, 2011. From tiger tonpanda: Animal head detection. IEEE T. Image Process., 20(6): 1696-1708J. Clerk Maxwell, A Treatise on Electricity and Magnetism, 3rd ed., vol. 2. Oxford: Clarendon, 1892, pp.68–73.

[2] Cui, J., F. Wen and X. Tang, 2008. Real time Google and live image search re-ranking. Proceedings of the 16th ACM International Conference on Multimedia, NY, USA, pp: 729-732

[3] Dalal, N. and B. Triggs, 2005. Histograms of oriented gradients for human detection. Proc. CVPR, 1: 886-893. Felzenszwalb, P. and D. Huttenlocher, 2005. Pictorial structures for object recognition. Int. J. Comput. Vis., 61(1): 153-190.

[4] Hu, J. and A. Bagga, 2003. Functionality based web image categorization. Proceeding of International World Wide Web Conference, pp: 118.

[5] Jones, M.J. and J.M. Rehg, 1999. Statistical color models with applications to skin detection. Proceeding of IEEE Conference on Computer Vision Pattern Recognition, pp: 81-96. Kun, X. and P. Xu, 2009.

[6] **A** human detection algorithm based on improved adaboost. Electr. Techn.(Shanghai), 11: 69-71.Lowe

[7] D.G., 1999. Object recognition from local scaleinvariant features. Proc. ICCV, 2: 1150-1157. Mikolajczyk, K., C. Schmid and A. Zisserman, 2004.

[8] Human detection based on a probabilistic assembly of robust part detectors. Proc. ECCV, 1: 69-82. Ru-jin, H., *et al.*, 2010.

[9] Adaboost for pedestrian detection based on multifeature. J. Jilin U. Sci. Edit., 48(3): 449-455.518. Schiele, B., 2005. Pedestrian detection in crowded scenes. Proc. CVPR, 1: 878-885.

[10] Zhu, Q., M.C. Yeh, K.T. Cheng and S. Avidan, 2006. Fast human detection using a cascade of histograms of oriented gradients. Proc. CVPR, 1491: 1498.

[11] P. Viola and M. J. Jones, "Robust real-time face detection," *Int. J. Comput. Vis.*, vol. 57, no. 2, pp. 137–154, May 2004.

[12] C. Cortes and V. Vapnik, "Support-vector networks," *Machine Learning*, vol. 20,no. 3, pp. 273–297, 1995.

[13] Research Journal of Applied Sciences, Engineering and Technology Organization, 2012 Corresponding Author: R. Rakesh, School of Computing, SASTRA University, India 5344 Human and Animal Detection Systems-[Hads] for Unauthorized Zone

[14] Kun, X. and P. Xu, 2009. A human detection algorithm based on improved adaboost. Electr. Techn.(Shanghai), 11

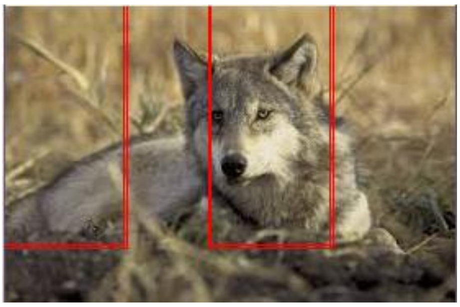
[15] 69-71.Lowe, D.G., 1999. Object recognition from local scaleinvariant features. Proc. ICCV, 2: 1150-1157.

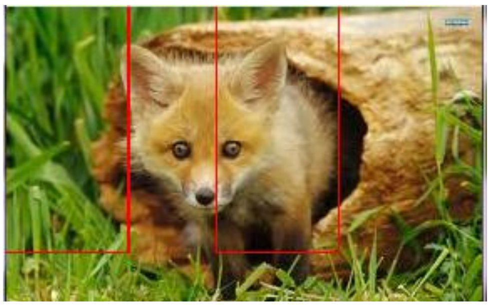
[16] B. Leibe, E. Seemann, and B. Schiele, "Pedestrian detection in Crowded scenes," in Proc. CVPR, 2005, vol. 1, pp. 878–885.

Mr. M.B. Rangdal has 6.2 years of experience teaching undergraduates. He is now pursuing ME degree, in Computer Engineering at VPCOE, BARAMATI.He has attended number of workshops on Robotics, FE/.SE Syllabus Setting, Research paper writing, Latex, Computer Vision , Akash Tablet for Education conducted by IIT, Bombay remote center at VPCOE, Baramati.

Dinesh Bhagwan Hanchate Birth Place:- Solapur, B.E. Computer from Walchand College of Engineering, Sangli (1995), Lecturer in Gangamai College Of Engineering,Dhule (1995-96), Lecturer in S.S.V.P.S.s B.S.D. College Of Engineering,Dhule In Computer & IT deptt (1996-2005), M.Tech. Computer from Dr. Babasaheb Ambedkar Technological University,

Lonere(2002-05), Currently Asst. Prof. in PG Computer Engineering, former H.O.D. (Computer IT) in Vidya pratishthans College Of Engineering, Baramati , currently doing research (SGGSs Institute of Technology and Engg, Nanded a_liated to SRTMU,Nanded) under the guidance of Dr. Bichkar R.S. ,G.H. Raisonis College Of Engineering and Management,Wagholi,,Pune.





ISSN: 2321-8169
178 – 183

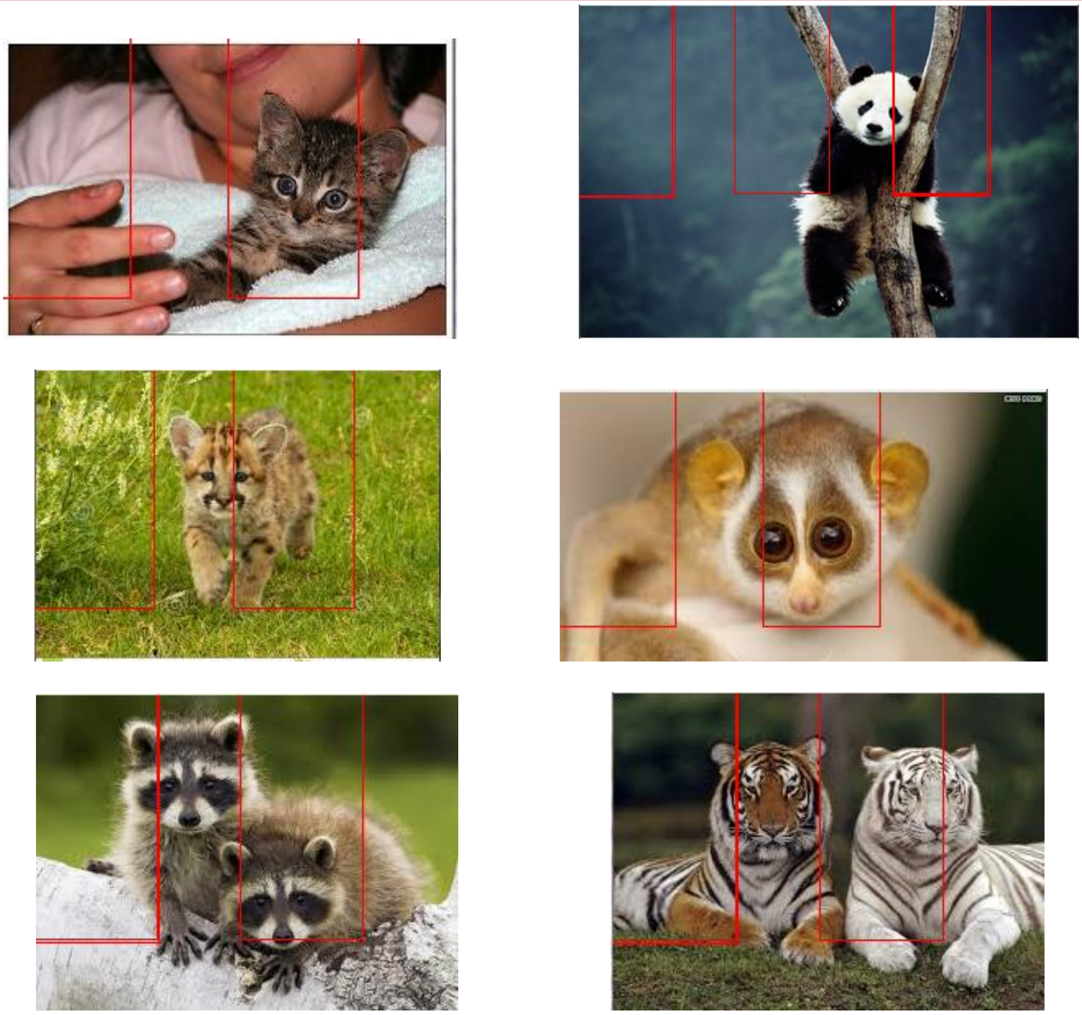


Figure 9: Exprimental Results