ISSN: 2277-9655
Impact Factor: 5.164
CODEN: IJESS7

IJESRT

INTERNATIONAL JOURNAL OF ENGINEERING SCIENCES & RESEARCH TECHNOLOGY

ANIMAL ACTION RECOGNITION: ANALYSIS OF VARIOUS APPROACHES

Ginet George[1], Anjali Namdev[*2] & SarangaPani Sarma[3]
[1]M.Tech, Department of Computer Science Engineering, SITM, Rewari (Haryana)
[*2]Assistant Professor, Department of Computer Science Engineering, SITM, Rewari (Haryana)

DOI: 10.5281/zenodo.1219671

ABSTRACT
Activity recognition and tracking of animals in video sequences is the main objective of this work. Each video is divided into 200 frames. Histogram of Oriented Optical flow (HOOF) features of size 32 bins are calculated for each of these frames and merged for a particular action .K-means clustering technique is being applied on these merged feature of size 2000x32 to find out 100 cluster centers each of size 32. Distance calculation of HOOF feature for each video sequence of 200 frames with cluster centers is being carried out to find the closest cluster centre for a particular action. Feature vector of size 200 is generated for each video sequence based on the minimum total distance of all the frames of a video with the cluster centers. This feature vector is used to train Support vector machine and K-Nearest neighbors algorithm for recognition of four different actions Running, walking, jumping and resting for a particular animal and the result of various algorithms involved here are compared based on their performance.

KEYWORDS: Animal detection, Activity Recognition, HOOF Feature, SVM, KNN

I. INTRODUCTION
Technology has upgraded immensely, such a technology is object tracking in video. Numerous research works have been pulled into the technology, which proves to be an important implementation with much more advancements researches in computer vision. Video surveillance has availed various applications such as animal tracking, video communication, brain-computer interaction, safety and supervision systems, traffic monitoring and controlling. In a cluttered scenario, confrontations arise in tracking the target in video surveillance activities. Object detecting, classifying, tracking and identifying the various activities are the primary objectives that needs to be accomplished while video surveillance [1]. In the starting age of animal activity analysis, some video sequences of animals are recorded for a period of time, and then a human observe the video and records the several activities of the animals manually. And this is a time consuming task and lots of hard work is required to watch the videos and records the activities. To overcome this situation, many researchers have proposed automated video processing methods to records the activities of animals. In this paper, we are focusing on some existing techniques and our proposed method.

II. LITERATURE REVIEW
Very few works have been reported to detect animal activities in various scenarios. In previous works, mostly manual observations and wearable sensors are utilized to track and record animal activities. Collar-worn accelerometers have been used to identify a variable range of activities: commercial gadgets, for example, Whistle or FitBark perform fundamental activity level recognition, like resting versus moving and so on. While the work by Ladha et al. is able distinguish between a more extensive range of behaviors on dogs (14 exercises and 2 postures)[2]. Samarasinghe et al. [3], collected the data via direct observation and total 840 minutes of observation were recorded to analyze various activities. In the work of Biolatti et al. [4], five observer were involved for data collection and watching the videos to analyze various activities of captive tigers. Godsk [5] Studies on the subject of recognizing cow activities. As a source of dataset, they have used satellite based position data. Engelberg et al.[6] uses GIS data to analyze the dog walking activity with adolescents' moderate-to-vigorous physical activity (MVPA) and BMI, and analyze the correlations of various dogs. They have found that, walkers had 7–8% more minutes/day of MVPA than non-dog walkers, and correlates of dog walking were found at multiple levels of influence. Soltis et al. [7], used wearable accelerometers to identify elephant activity levels

ISSN: 2277-9655
Impact Factor: 5.164
CODEN: IJESS7

and body orientation. They have considered 6 elephants to analyze the activity and body orientation. M Zeppelzauer [8] proposed an automated technique in his research work for the identification and tracking of elephants in wild life video. He prepare a shading model from videos to detect elephants in wild life videos. Z He et al. [9] developed a camera sensor called Deer Cam. Important video and sensor data about activities are collected by mounting the sensor in animal body. Further, the data are used on various scientific research to analyze activities. In the work of Y Iwashita et al. [10], they have introduces the first-person animal activity recognition, A camera is attached with the animal to monitor and understand the behaviors of animal in absence of human. S Pedersen et al. [11], developed a novel activity sensing system which measures activity of domestic animals. Passive infrared detectors (PID) and a specially designed analogue signal processing interface is used for collecting and analyzing data to recognize activities of animals.

III. PROPOSED METHOD

Algorithm 1: Object Activity Recognition
Input: Video displaying different animal activity.
Output: Activity in the video is recognized.

Step1: Capture and categorized video displaying different activity of a tiger. A video must contain only one activity.
Step2: Divide the video into 200 frames and calculate optical flow for each of the frames as explained in section 3.1.
Step3: Extract 32 bin size HOOF feature for each frame as explained in section3.2.
Step4: Combine all the HOOF features for a particular video to generate 32x200 features.
Step5:10 videos for every action is taken for training purpose. Merge all the HOOF features for a particular action to generate TOTAL_HOOF of size 32x200x10. As the number of activities are four, four different TOTAL_HOOF are generated.
Step6: Perform K-Means Clustering on TOTAL_HOOF to find out 100 cluster centre each of size 32 for each action.
Step7: Euclidian Distanceis calculated between cluster centers of TOTAL_HOOF and HOOF of each video to find out the nearest cluster centre of size 32 and a feature vector for each action is generated and lable it.
Step8: SVM [12] and KNN [13] classifiers are usedfor training and testing.

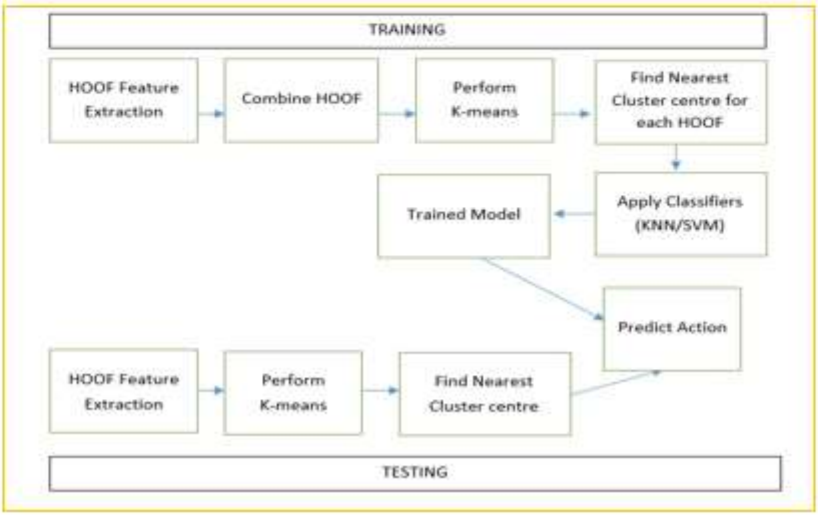

Fig. 1.Proposed Methodology

IV. OPTICAL FLOW

The term “optical flow” refers to a visual phenomenon that we experience every day. Optical flow is the apparent visual motion that we experience as we move through the world. Suppose we are sitting in a car or a train, and are looking out the window. We will see trees, the ground, buildings, etc., appear to move backwards. This motion is optical flow. The goal of optical flow[14] estimation is to compute an approximation to the motion field from time-varying image intensity. Fig. 2 shows optical flow of a **Frame.**

ISSN: 2277-9655
Impact Factor: 5.164
CODEN: IJESS7

Fig. 2. Optical Flow of one frame

HISTOGRAM OF ORIENTED OPTICAL FLOW (HOOF):

In our purpose we are extracting the HOOF features [15] from the videos. In Histogram of Oriented Optical Flow (HOOF) feature, optical flow is calculated for each frame of the video and then HOOF is calculated for each block b in the frame. With respect to the primary angle from the horizontal axis, each flow vector is binned and weighted according to its magnitude. Here, primary angle refers to the smallest signed angle of optical vector from horizontal axis. All optical flow vectors, v is $[x,y]^t$ with direction $\theta = \tan^{-1}(y/x)$ in the range

$$-\frac{\pi}{2}+\pi\frac{b-1}{2} \leq -\frac{\pi}{2}+\pi\frac{b}{B} \tag{1}$$

Where $b = \sqrt{x^2+y^2}$, $1 \leq b \geq B$, B = total no. of bins, b = bin

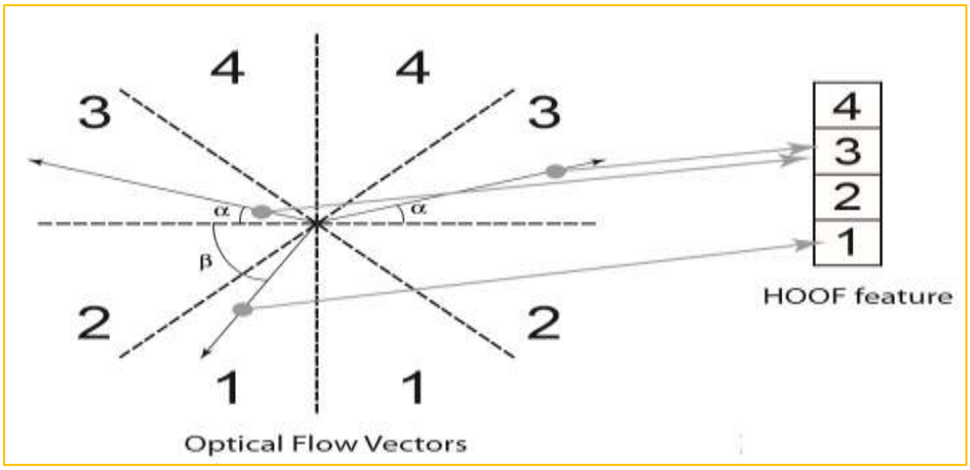

Fig.3 Four Bin Histogram formation

Finally the histogram is normalized up to 1 which makes the histogram representation scale invariant. Hence, the histogram will not be affected by the distance between object and the camera i.e. the object is at far distance or near the camera. Fig 3 shows a four bin histogram formation. Histogram representation is direction independent of motion i.e. the histogram will be same whether the object is moving from right to left or left to right direction because of binning with respect to primary angle. If background is stationary, there will be no optical flow, hence background subtraction and object segmentation is not required in extraction of HOOF features. Small noisy optical flow will have small effect on the histogram as the contribution of optical vector is

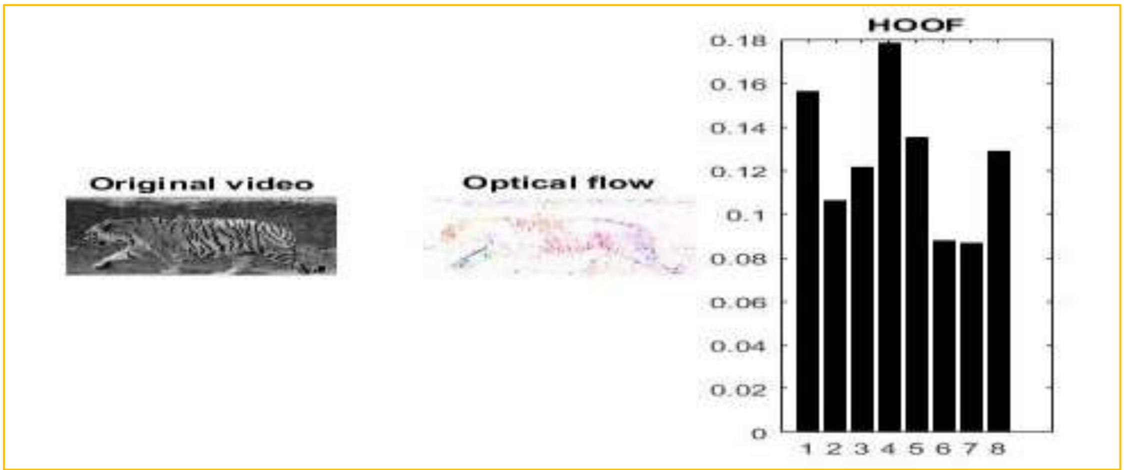

Fig. 4.Optical Flow and 8 bin Histogram projection

V. FEATURE VECTOR GENERATION AND CLASSIFICATION

10 videos for every action is taken for training purpose. Merge all the HOOF features for a particular action to generate TOTAL_HOOF of size 32x2000. As the numbers of activities are four, four different TOTAL_HOOF is generated. Perform K-Means Clustering on TOTAL_HOOF to find out 100 cluster centre each of size 32 for each action. Euclidian Distance is calculated between cluster centers of TOTAL_HOOF and HOOF of each video to find out the nearest cluster centre of size 32 and a feature vector for each action is generated and labeled.
As the number of actions considered is four, the number of classes of this classification problem will be four. For solving the present classification problem two different classifiers namely: Support Vector Machine and K-NN classifier are used on the feature vector generated to perform the task of classification.

VI. DATASET PREPARATION

As the benchmark dataset for this particular task is not available as such, we have prepared a novel dataset for this particular experiment. one particular animal species is considered here. the presented datasets are used throughout this work to evaluate the proposed approaches. the dataset consists of animated videos as well as realistic videos with cluttered background. the dataset contains 4 actions of tiger. i.e. walking, running, jumping, resting. each action contains 10 videos of 200 frames. to deal with real life scenario, the dataset collection is done from various sources like youtube and google videos and various websites like shutterstock.com. fig.5 shows three sample video frames of four different actions: walking, running, jumping and resting.

ISSN: 2277-9655
Impact Factor: 5.164
CODEN: IJESS7

Fig. 5 three sample video frames for each of the four different actions.

VII. EXPERIMENTAL RESULT, COMPARISON AND ANALYSIS

In this section, an evaluation, comparison, and analysis of our action recognition framework is done. Extraction of Histogram of Oriented Optical Flow features from each video. Merging of all the features of a particular action is done and K-Means Clustering is performed to find the cluster centers. Closest cluster center for each HOOF feature sequence of each video is calculated. SVM and K-NN classifiers are used for training and testing. 1600 frames are used for training a particular action and 400 frames are used for testing the performance. For SVM, three different kernels are applied, i.e. cubic, quadratic and linear, and among them cubic SVM performs well with 87.5% accuracy. Fine KNN and weighted KNN are used, which shows 92.5% and 75% accuracy respectively. Table 1 shows a comparison between various approaches.

Table 1.Comparison of various Approaches and HOOF feature

Feature	Classifier	Accuracy
HOOF feature	Fine KNN	92.5%
HOOF feature	Cubic SVM	87.5%
HOOF feature	Quadratic SVM	82.5%
HOOF feature	Linear SVM	75.0%
HOOF feature	Weighted KNN	75.0%

Fig. 6.Confusion matrix for Fine KNN

A confusion matrix represents the recognition accuracy of a classification model over a set of labeled test data. The Confusion matrix corresponding to the best performing classifier Fine KNN is shown in Fig.6

VIII. CONCLUSION AND FUTURE WORK

This research work presents a new efficient classification approach for animal activity recognition. Novel dataset for animal action recognition is prepared which can be used for further research. This research investigates performance of SVM and KNN with different kernels with HOOF feature for animal activity recognition. Results shows that the Fine KNN outperforms the rest kernels in the recognition task of videos taken into consideration. Though the research investigates different actions by same animal species, detecting multiple animal activities in the same video is out of the scope of this research work, which can be investigated in future.

IX. REFERENCES

[1] Ojha, S., &Sakhare, S. (2015, January). Image processing techniques for object tracking in video surveillance-a survey. In Pervasive Computing (ICPC), 2015 International Conference on (pp. 1-6). IEEE.

[2] Ladha, C., Hammerla, N., Hughes, E., Olivier, P., &Ploetz, T. (2013, September). Dog’s life: wearable activity recognition for dogs. In Proceedings of the 2013 ACM international joint conference on Pervasive and ubiquitous computing (pp. 415-418).ACM.

[3] Samarasinghe, W. M. P., &Ahamed, A. R. (2016). A Preliminary study on Activity Budgets of Asian Elephant (Elephasmaximus Linn.) at Elephant Orphanage. *Bull. Env. Pharmacol. Life Sci*, *5*, 47-50.

[4] Biolatti, C., Modesto, P., Dezzutto, D., Pera, F., Tarantola, M., Gennero, M. S., ...&Acutis, P. L. (2016). Behavioural analysis of captive tigers (Pantheratigris): A water pool makes the difference. Applied Animal Behaviour Science, 174, 173-180.

[5] Godsk, T. (2011). Methods and Software Architecture for Activity Recognition from Position Data.

[6] Engelberg, J. K., Carlson, J. A., Conway, T. L., Cain, K. L., Saelens, B. E., Glanz, K., ... &Sallis, J. F. (2016). Dog walking among adolescents: Correlates and contribution to physical activity. *Preventive medicine*, *82*, 65-72.

[7] Soltis, J., King, L., Vollrath, F., & Douglas-Hamilton, I. (2016). Accelerometers and simple algorithms identify activity budgets and body orientation in African elephants Loxodontaafricana. *Endangered Species Research*, *31*, 1-12.

[8] Zeppelzauer, M. (2013). Automated detection of elephants in wildlife video.EURASIP journal on image and video processing, 2013(1), 1.

[9] He, Z., Eggert, J., Cheng, W., Zhao, X., Millspaugh, J., Moll, R., ...&Sartwell, J. (2008). Energy-aware portable video communication system design for wildlife activity monitoring. *IEEE Circuits and Systems Magazine*, *8*(2).

[10] Iwashita, Y., Takamine, A., Kurazume, R., &Ryoo, M. S. (2014, August). First-person animal activity recognition from egocentric videos. In *Pattern Recognition (ICPR), 2014 22nd International Conference on* (pp. 4310-4315). IEEE.

[11] Pedersen, S., & Pedersen, C. B. (1995). Animal activity measured by infrared detectors. *Journal of Agricultural Engineering Research*, *61*(4), 239-246.

[12] Cortes, C., &Vapnik, V. (1995). Support-vector networks. *Machine learning*, *20*(3), 273-297.

[13] Keller, J. M., Gray, M. R., & Givens, J. A. (1985). A fuzzy k-nearest neighbor algorithm. *IEEE transactions on systems, man, and cybernetics*, (4), 580-585.

[14] Horn, B. K., &Schunck, B. G. (1981). Determining optical flow. *Artificial intelligence*, *17*(1-3), 185-203.

[15] Chaudhry, R., Ravichandran, A., Hager, G., & Vidal, R. (2009, June). Histograms of oriented optical flow and binet-cauchy kernels on nonlinear dynamical systems for the recognition of human actions. In Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on (pp. 1932-1939). IEEE.

CITE AN ARTICLE

George, G., Namdev, A., & Sarma, S. (n.d.). ANIMAL ACTION RECOGNITION: ANALYSIS OF VARIOUS APPROACHES. *INTERNATIONAL JOURNAL OF ENGINEERING SCIENCES & RESEARCH TECHNOLOGY*, *7*(4), 548-554.