Speeded Human Action Recognition Based on Action Snippets Algorithm and Neighborhood Component Analysis

Sanjay T.Gandhe, Pravin A.Dhulekar

Abstract*: The human action recognition has large number of applications in various domains however its applicability in many real time scenarios has been restricted due to the speed of action recognition algorithms. Therefore there is great need to develop the faster action recognition algorithm. To address this constraint we present a hybrid algorithm for speeded human action recognition (SHAR) based on action snippets algorithm and neighborhood component analysis. The action snippet algorithm results in filtering the key frames and eliminating less informative frames. Moreover the feature selection algorithm at further stage of recognition helps to trimmed down the length of feature vector massively. These two algorithms are integrated with standard pair of feature extraction and classification at different stages to accomplish the goal of faster recognition without compromising the recognition accuracy. The performance of proposed algorithm has been evaluated on weizmann dataset. The proposed methodology provides overall recognition accuracy of 90% with four times faster classification over standard approach of feature extraction followed by classification.

Index Terms: speeded human action recognition (SHAR), action snippets algorithm and neighborhood component analysis.

I. INTRODUCTION

In last few years the human action recognition fascinated significant attention among several researchers because of its large number of applications in various domains that includes human–computer interaction, biometrics, telehealth, and video indexing [1]. The human action recognition is the process of identifying the name of action based on the various body movements performed by human in their day to day life. In short, giving suitable name to the action that can be understood by any average human [2]. Based on learnt knowledge the system will be able to assign the accurate name to the action. Overall the labeling accuracy depends on the training phase which mainly depends on two steps; feature extraction and classification. The feature extraction is the process to extract the most meaningful features that can best way describe the contents of input database. The extracted features are fed to classifier which classifies the action based on learnt knowledge during the training phase [3]. The optimized training results in least classification error and least over fitting issues. So all is depends on the proper selection of feature extraction and classification method because they are complimentary to each other which means if the classifier is weak then extracted features should be enough strong to obtain the desired recognition accuracy or in case extracted features lags behind to describe the person movement then classifier should have tweaking factor which can compensate the role of feature extractor. Overall these methods should work hand in hand to provide the best results for the task of human action recognition [4].In this paper, we have focused on how to minimize the feature extraction overhead, classifiers fitting and prediction period to provide speeded human action recognition. To achieve the same we have divided our work in two parts, in first we have trimmed down the input data at the beginning itself in large extent using action snippets algorithm which reduces the feature extraction overhead enormously. At second stage we are applying the feature selection algorithm over extracted features to filter out most valuable and meaningful features by eliminating the irrelevant features. This two stage work help classifier to have most appropriate data for training and testing that not only reduces classifiers fitting and prediction period but also helps in minimizing the classification and over-fitting errors.The paper has been arranged in five sections, first section introduces the outline of work, in second related work has been reviewed. The third and fourth section covers two stages proposed methodology and in fifth section result analysis has been presented to draft the most meaningful conclusions in sixth section which can help further researchers to carry out their work in appropriate direction.

Revised Manuscript Received on June 07, 2019.

Dr Sanjay T. Gandhe, Professor, Department of Electronics and Telecommunication Engineering, Sandip Institute of Technology and Research Centre, Nashik, Savitribai Phule Pune University, Pune, India

Pravin A. Dhulekar, Research Scholar, Department of Electronics and Telecommunication Engineering, Sandip Institute of Technology and Research Centre, Nashik, Savitribai Phule Pune University, Pune, India.

II. RELATED WORK

As discussed in the earlier section that this work focuses on development of faster human action recognition technique in which major role is of appropriate features. As perfect features can help to improve the action recognition performance in the similar way irrelevant features can deteriorate the system performance in terms of wastage of fitting and prediction period as well as reduced recognition accuracy. In the view of same the review has been carried out for two purposes; first to select the most important frames among the recorded video called as snippets and second to select the most important features from these snippets.

A. Frame Selection Methods

Always there is a question about how many frames are needed for accurate recognition of human actions.

Generally the videos are recorded at the rate of 24fps to 30fps. So even 5 seconds video results in around 120 to 150 frames but all they are really needed or whether they all content the significant information. The answer is surely not positive because in many scenarios even couples of frames are sufficient to describe the action accurately. Therefore here the need of frame selection method has been raised. Few works has been carried out in that direction. The overview of these methods has been detailed as:Efros et al. introduced optical flow based motion descriptor in spatio-temporal volume for every stationary human skeleton. The classifier used to label at every frame level where around 50 frames has been used for action recognition. In this work noisy optical flow measurement was also handled by considering the optical flow as spatial pattern. This was the first attempt to identify the action at frame level [5].Jhuang et al. developed biologically-motivated system for the human action recognition from video sequences. Author shown that use of sparse features filtered from feature selection methods gives better results over dense one. With least number of features author has achieved higher recognition accuracy [6].Konrad Schindler and Luc Van Gool presented an algorithm that recognizes the human action from just less than 10 frames. Author evaluated the algorithm on standard datasets and achieved the accuracy around 90%. For this work author used local shape and optic flow features. The results has also been compared with accuracy obtained using complete video sequence and proved that even 5-7 frames can also provide the similar accuracy. Our approach of snippet algorithm has been inspired from this work [7].Niebles and Li developed the model able to classify human actions at frame by-frame level. Author has extracted static and dynamic interest points by representing video sequence as bundle of spatial and spatial-temporal features. Combined use of static and dynamic features gives rich representation of action than individual feature type [8].Laptev and Lindeberg represented and recognized the motion patterns contents in video from set of local space-time descriptors. In this work classification is executed at frame level. Various image descriptors were evaluated for different human activity recognition. Author has highlighted the significance of local position dependent histograms over spatio-temporal jets and position independent histograms [9]. Ali et al. adopted chaotic systems based model to analyze the non linear dynamics of human actions. Trajectories of reference joints are used to extract invariant features that reconstruct phase space which represent the dynamical system. The combination of these invariants over all the trajectories has been used to represent the action. Labeling is done using exemplars with KNN classifier at sequence level [10].

B. *Feature Selection Methods*

Feature selection is the technique of selecting a small subset from a given set of features by eliminating irrelevant and redundant features. Common approaches to feature selection include stepwise regression, sequential feature selection, and regularizationIn recent years several methods has been proposed by researchers for the task of feature selection. Here we will take review of some noteworthy contributions.R. Gilad-Bachrach introduced margin based feature selection method to quantify the valuable feature sets. Margins have been used to provide theoretical generalization bound and to solve the classification problems in multi class applications. The applicability of the algorithm has been proven by comparing it with Relief algorithm based on ratio of accuracy with number of features [11]. Y. Sun developed specifically for the data with large number of irrelevant features. Using local learning arbitrarily complex nonlinear problem has been divided into locally linear sets. The work is based on support vector machine and numerical analysis techniques. Author has claimed that proposed algorithm can process several thousand features in minutes without compromising the accuracy which is independent of the increase in number of irrelevant featurers. Effectiveness of the algorithm has tested over non conventional datasets such as spiral, thyroid, diabetics, splice, breast cancer and DLBCL dataset [12]. B. Chen considered multimodally distributed data for feature selection and presented the large margin feature weighting method for k-nearest neighbor classifier. By reducing the cost function algorithm learns the feature weighting factors derived for separating different classes by large local margins and bounding together the closely related points under the same class. This algorithm provides efficient solution based on linear programming. The proposed work has been tested on UCI and microarray data sets [13].

III. PROPOSED METHODOLOGY

In this paper we are proposing the novel methodology for faster recognition of human actions based on hybrid combination of various methods at various levels that is action snippet algorithm for frame selection, histogram of oriented gradients (HOG) for feature extraction, neighborhood component analysis (NCA) method for feature selection and at the last support vector machine (SVM) for action classification. In this work we have applied best among the available methods for each specific task. To extract action features various extraction technique are available like spatio temporal interest points (STIP), motion history volumes (MHV), motion history images (MHI) and motion energy image [14]. The earlier research pointed STIP as state of art in the domain however [15] shown that HOG is more effective over STIP in terms of recognition accuracy and response time. Therefore for feature extraction we are using HOG. In the similar way among various supervised learning method, the SVM classifier has been selected based on its least training and testing time as well as higher classification accuracy [15]. The figure 3.1 shows the complete flow diagram of proposed methodology. After acquisition of action videos, the first and foremost step is to obtain segmented frames from this recorded videos that can be easily achieved using the method of temporal segmentation. The video after segmentation results in large number of frames, obviously each of these frames doesn’t contain the meaningful information. So here the task is to filter out the most informative frames and to skip frames with least information. This task is fulfilled by the use of action snippet algorithm which is explained in further subsections. Snippets undergoes preprocessing task where resizing and gray scale conversion operation is applied to have uniform datasets.

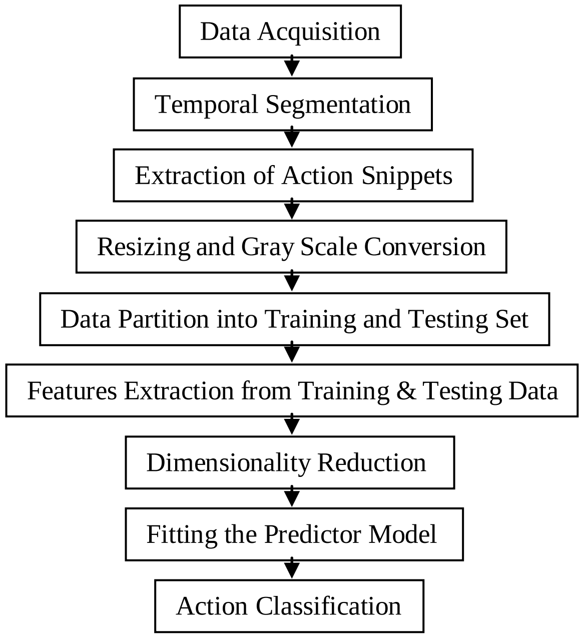


Fig 3.1 Flow diagram of proposed methodology

The preprocessed snippets are further partitioned into training and testing set. There are different ways of partitioning the data on the basis of various cross validation techniques such as leave one-out cross-validation, 4-fold cross-validation, 5-fold cross-validation. As per the dataset here we are using leave one-out cross-validation technique. Features are extracted individually from this training and testing datasets. As specified earlier we are using HOG for extraction of features. The extracted features are huge in volumes which contain several irrelevant features that need to be filtered out so that overall classification time can be reduced. For the same we are going for dimensionality reduction that can be achieved using suitable feature selection algorithm which is detailed in third subsection of section IV. As only these selective features are used for training the classifier so overall fitting and prediction time get reduced significantly which makes classification process faster. In the last step the fitted model was used to accurately predict the class of action.

IV. SPEEDED HUMAN ACTION RECOGNITION

The speed of action recognition can be increased only if action can be recognized on the basis of least amount of data. This work continuously focuses on reducing the amount of data required for action recognition without compromising recognition accuracy. To achieve the same two novel algorithms one is action snippet algorithm and other is neighbor component analysis algorithm has been hybridized along with standard process of feature extraction followed by classification.

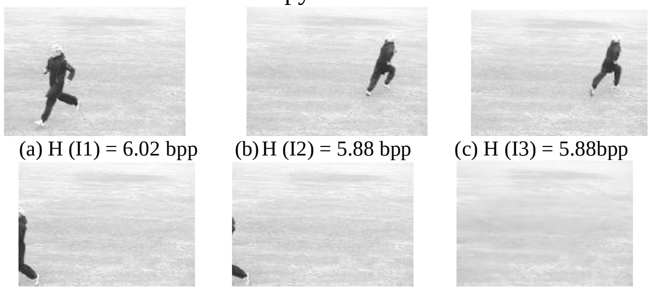
A. Action Snippets Algorithm

Snippets are defined as the key frames extracted from video that holds significant amount of data. Action snippets are the frames with substantial amount of information related to specific action [7]. The obvious purpose of snippets is to reduce the number of required frames that leads to faster and optimized human action recognition. The action snippets algorithm has been developed to eliminate subject less frames, frames in which subject is partially visible, frames with repetitive information specifically occurs in case of cyclic/periodic actions.

The proposed action snippet algorithm is based on computing the entropy of the image. The entropy is measure of degree of randomness used to characterize the image. The entropy or average information of an image is calculated from the histogram of the image as histogram shows the various grey level probabilities in the image. The equation for the entropy (H) of image (I) is given in (1).

$$H(I) = - \text{sum}(p.*\log2(p)) \quad (1)$$

Where p is the normalized histogram counts obtained from image histogram.Once the entropy of all frames in the video sequence has been computed then frames are aligned as per the descending value of entropy. From the figure 4.1 it can be observed that the subject less frames (f) and frames with partial appearance of subject (d & e) has a lowest entropy value. On the basis of which they can be easily eliminated. Also repetitive frames (b & c) can be avoided based on similarities in their entropies. The key frame (a) with substantial amount of information always ranks at initial order in terms of their highest entropy value. Here the number of frames to be filtered out as key frames depends on the user defined threshold for entropy value.



(a) H (I1) = 6.02 bpp (b) H (I2) = 5.88 bpp (c) H (I3) = 5.88bpp
(d) H (I4) = 3.72 bpp (e) H (I5) = 3.67 bpp (f) H (I6) = 1.43 bpp

Fig 4.1 Entropy of the images

B. Feature extraction using HOG

The histogram of oriented gradients (HOG) feature descriptor method is based on computation of normalized local histograms of image gradient orientations. The central concept is person's appearance and shape can be defined accurately from distributions of local intensity gradients or from directions of edges [16]. The stepwise operations involved in HOG are given below:

1. In the preprocessing stage entire image or patch is preprocessed. In case of patch the desired part of image is cropped first and then resized.
2. In this step, the gradient image is calculated by computing magnitude of gradient (g) and direction of gradient (ϴ) using following equation (2) & (3):

$$g = \sqrt{gx^2 + gy^2} \quad (2)$$

$$\theta = \tan^{-1}\frac{gy}{gx} \quad (3)$$

Here the gx and gy are the x and y derivatives of the image which are computed by convolving the image with horizontal (fx) and vertical (fy) kernels such as [17]:

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

$$f_x \Rightarrow \frac{1}{3}\begin{bmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{bmatrix} \qquad f_y \Rightarrow \frac{1}{3}\begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ -1 & -1 & -1 \end{bmatrix}$$

When there is no change in the intensity values of image with respect to y the vertical gradient gy will be zero where as horizontal gradient gx shows values where there is change in intensity. This clearly indicates the applicability of gradient to capture the regions where abrupt change in intensity occurs.

3. In this step the histogram of gradient is computed by fixing the block size and cell size. The cell size is inversely proportional to length of feature that is when cell size is small the features length is large. Though large number of features contributes to better accuracy but they results in large computation time therefore the cell size should be optimistic.

The figure 4.2 shows the impact of cell size on resultant gradients [15].

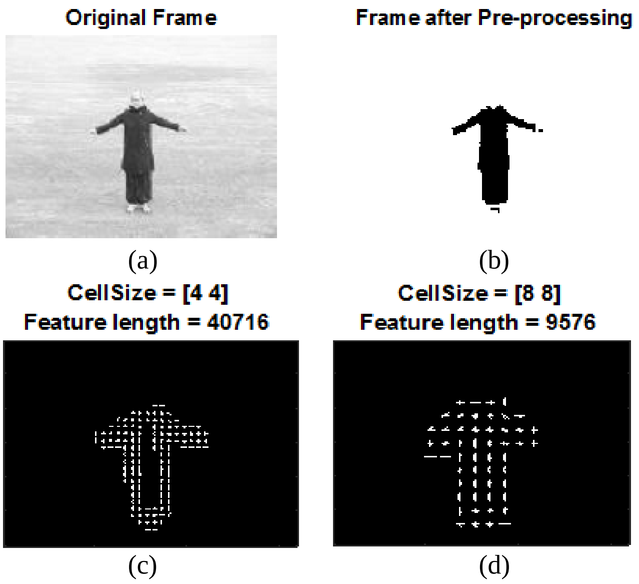


Fig 4.2 (a) Original frame (b) Preprocessed frame (c) resultant gradients with 4x4 cell size (d) resultant gradients with 8x8 cell size [15]

4. The gradients of the image are always sensitive to the lighting. To obtain the performance of HOG descriptor independent of lighting we go for block normalization as per the below equation.

$$\boldsymbol{v} \rightarrow \frac{\boldsymbol{v}}{\sqrt{\|\boldsymbol{v}\|_2^2}} \qquad (4)$$

Where v is the feature vectors

5. In the last step the final feature vector for the entire image is calculated by concatenating all feature vectors in one large vector.

C. Feature selection using NCA

The feature vectors dimensionality reduction can be obtained either by feature transformation method or by feature selection method. Feature selection method chooses a subset of inputs from many whereas Feature transformation method transforms many inputs into fewer inputs such as principal component analysis (PCA), linear discriminant analysis (LDA), or singular value decomposition (SVD) etc. Feature selection method retain original form of features however in feature transformation method may lose the features in the original space because of which interpretation may be different or very difficult. Also Feature selection can be done with simpler hypothesis that need very simple model and hence more interpretable compare to feature transformation. By considering all these parameters we have chosen feature selection method over feature transformation.To handle high-dimensional data the number of algorithms has been successfully developed for feature selection such as Large margin feature weighting (LMFW), Iterative Search Margin Based Algorithm (Simba), Local learning based feature selection (FSSun) and Neighborhood component feature selection (NCFS). Among all these techniques neighborhood component feature selection (NCFS) shows outperformance in terms of number of optimal features and minimum classification error [18]. In addition, the NCFS is almost insensitive to the increase in the number of irrelevant features and performs better than Simba, LMFW and FSSun in most cases.NCFS technique is based on Neighborhood component analysis (NCA) which is a powerful feature selection technique that can handle very high-dimensional datasets. It's nearest neighbor-based feature weighting algorithm, which learns a feature weighting vector by maximizing the expected leave-one-out classification accuracy with a regularization term.

Let T = {(x1, y1), . . . , (xi, yi), . . . , (xN, yN)} be a set of training samples,

where xi = d-dimensional feature vector.

yi ∈ {1, . . . ,C} = corresponding class label.

N = number of samples.

Here the objective is to compute a weighting vector w that lends itself to select the feature subset optimizing nearest neighbor classification. In terms of the weighting vector w, we denote the weighted distance between two samples xi and xj by equation (5).

$$D_w(\mathrm{x}_i, \mathrm{x}_j) = \sum_{l=1}^{d} \mathrm{w}_l^2 (\mathrm{x}_{il} - \mathrm{x}_{jl}) \qquad (5)$$

where w_l is a weight associated with *l*th feature & d is number of features. The feature weight can be estimated in equation (6).

$$\mathrm{w}_l = \frac{1}{k} \sum_{j=1}^{k} \max\left(0, 1 - \frac{\delta_{lj}^2}{\delta_l^2}\right) \qquad (6)$$

where l = 1,2,3………d

k = number of feature partitions

δ_{lj}^2 = variance of the observations in jth partition Along the *l*th dimension

δ_l^2 = global variance of all observations on the *l*th Feature. Feature can be regarded less relevant if the variance of observations in a partition is closer to the global variance. The overall contribution of feature selection algorithm in speeding up the action recognition has been analyzed in Vth section.

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

D. Action Classification

Out of several classification techniques selection of most appropriate classifier is the key challenge. Despite of number of selection criteria we have restricted our selection on the basis of major two factors one is classification accuracy and other as overall classification time. On the basis of different experimental analysis [15] we observed the performance of support vector machines (SVM) is in line with our requirement.SVM classifies data by finding the best hyperplane that separates all data points of one class from those of the other class. Support vectors are the data points that are closest to the separating hyperplane; these points are on the boundary of the slab [9].The figure 4.3 illustrates these definitions, with + indicating data points of type 1, and – indicating data points of type –1.

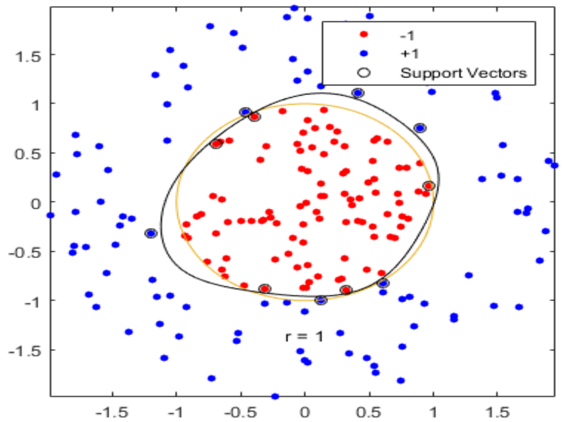


Fig 4.3 Support vectors for two class classification

The separating hyperplane in support vector machine is defined in equation (7)

$$w^T x + b = 0 \qquad (7)$$

Where w is a weight vector, x is input vector and b is bias. From this the equation for output class +1 and -1 are shown in equation (8)

$$w^T x + b \geq 0 \text{ for } yi = +1$$
$$w^T x + b < 0 \text{ for } yi = -1 \qquad (8)$$

where yi is the output class.

V. RESULT ANALYSIS

The proposed methodology has been evaluated on Weizmann dataset because of its inclusiveness towards major bottleneck issues in human action recognition such as execution rate, anthropometric variation, view invariance and camera motion. The Weizmann dataset is composed of 90 videos which includes the 10 actions (bend, jumping jack, jump forward, jump in place, run, gallop sideways, skip, walk, wave one hand and wave both hands) performed by 9 subjects. The sample frames of few actions have been shown in figure 5.1.

(a) (b) (c)

(d) (e) (f)

Fig 5.1 Weizmann frames of (a) Bend (b) Jack (c) Walk (d) Jump (e) Jump in place (f) Run

The experimental results have been analyzed on the basis of comparing the performance of proposed methodology with the existing outperforming method. The earlier study [15] shown that the combination of HOG with SVM gives best results over other action recognition methods. Therefore to augment the applicability of our proposed method we embed the action snippet algorithm and feature selection algorithm with HOG-SVM combination. The result has been compared on the basis of two major parameters recognition accuracy and recognition time. The proposed methodology has been abbreviated as SHAR (speeded human action recognition). The table I shows feature reduction (%) obtained using SHAR over HOG_SVM. Here we can observe the feature reduction for nine different cases which are based on leave one-out cross-validation that is out of nine subjects one subject is used for testing and eight for training. This process is repeated for all 9 permutations. The original training data size is 2240 x 432 where 2240 frames belongs to eight subjects and 280 frames in testing data size are belongs to one subject for each case. The average feature reduction for these nine cases is 74.62% which is significantly very high.

Table I: Feature Reduction using SHAR

Case No.	Training Data Size		Testing Data Size		Feature Reduction (%)
	HOG_ SVM	SHAR	HOG_ SVM	SHAR	
Case I	2240x432	2240x115	280x432	280x115	73.37%
Case II	2240x432	2240x120	280x432	280x120	72.23%
Case III	2240x432	2240x86	280x432	280x86	80.10%
Case IV	2240x432	2240x108	280x432	280x108	75.00%
Case V	2240x432	2240x108	280x432	280x108	75.00%
Case VI	2240x432	2240x109	280x432	280x109	74.77%
Case VII	2240x432	2240x122	280x432	280x122	71.76%
Case VIII	2240x432	2240x112	280x432	280x112	74.08%
Case IX	2240x432	2240x107	280x432	280x107	75.24%

Now we have to study the effect of these reduced features on recognition accuracy. The table II shows the comparative analysis of recognition accuracy obtained using HOG_SVM and proposed method (SHAR). The average recognition accuracy given by HOG_SVM was 85.51% whereas the SHAR has given 89.26% which shows around 3.75% improvement in recognition accuracy at video level.

Table II: Recognition accuracy (%) at the level of video

Case No.	HOG_ SVM	SHAR	Improved Accuracy (%)
Case I	82.14	88.35	6.21
Case II	86.07	89.37	3.3
Case III	91.42	93.56	2.14
Case IV	82.14	88.90	6.76
Case V	75.35	81.40	6.05
Case VI	89.64	92.36	2.72
Case VII	80.35	84.44	4.09
Case VIII	96.78	97.12	0.34
Case IX	85.71	87.90	2.19

The second most important performance metric is classification time which includes time taken by classifier for training and for prediction i.e. testing time. The table III depicts entire information related to time taken by HOG_SVM and SHAR.

Table III: Training, Prediction and Total classification time

Case No.	Training Time (Seconds)		Prediction Time (Seconds)		Total Classification Time (Seconds)	
	HOG_SVM	SHAR	HOG_SVM	SHAR	HOG_SVM	SHAR
Case I	3.90	1.03	0.18	0.04	4.08	1.07
Case II	4.36	1.21	0.19	0.05	4.55	1.26
Case III	4.37	0.86	0.19	0.03	4.56	0.89
Case IV	3.95	0.98	0.19	0.04	4.14	1.02
Case V	4.00	1.00	0.18	0.04	4.18	1.04
Case VI	4.38	1.10	0.18	0.04	4.56	1.14
Case VII	3.74	1.05	0.18	0.05	3.92	1.1
Case VIII	4.43	1.14	0.20	0.05	4.63	1.19
Case IX	4.40	1.08	0.19	0.04	4.59	1.12

The average training time for SHAR is 1.05 seconds which is three times less compared to that of taken by HOG_SVM i.e. 4.17 seconds. Similar performance has been recorded for prediction where SHAR takes 40 msec while HOG_SVM needs 180 msec.

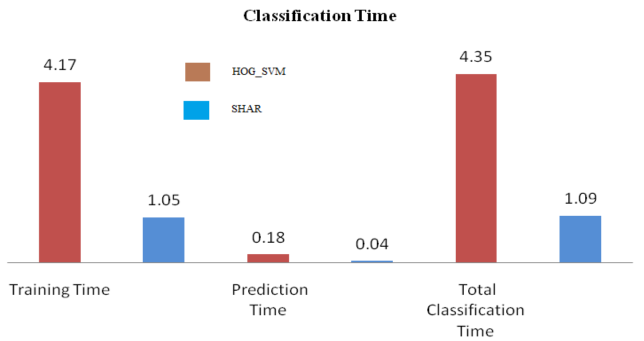


Fig 5.2 Average training, prediction and classification time

The bar plot in figure 5.2 shows comparative analysis between HOG_SVM and SHAR on the basis of average training time, average prediction time and average total classification time. the comparative The total classification time is the addition of training time and prediction time. Average total classification time required for HOG_SVM is 4.35 seconds whereas for SHAR it is recorded as only 1.09 seconds. Ultimately the SHAR provides around four times faster action recognition compared to HOG_SVM.

VI. CONCLUSION

This paper presents the method for faster human action recognition which is based on hybrid combination of action snippet algorithm and feature selection method with best pairing of feature extraction and classification method. The proposed method has been evaluated on standard dataset with various performance parameters. The results showed the significant improvement in recognition accuracy and reduction in overall classification time. This entire work strive this area towards next step of high speed action recognition that ultimately eliminate the limitations of human action recognition for various real time applications where speed is the utmost priority.

REFERENCES

1. Manoj Ramanathan, Wei-Yun Yau and Eam Khwang Teoh, "Human Action Recognition with Video Data: Research and Evaluation Challenges", IEEE TRANSACTIONS ON HUMAN-MACHINE SYSTEMS, 2168-2291, pp.1-14, 2014.
2. Lei Wang, Yangyang Xu, "Human Action Recognition by Learning Spatio-Temporal Features with Deep Neural Networks", IEEE Special Section On Cloud-based Robotic Systems for Intelligent Services, April 2018, PP.01-10.
3. Soumalya Sen, Moloy Dhar, "Implementation of Human Action Recognition using Image Parsing Techniques", Emerging Trends in Electronic Devices and Computational Techniques (EDCT), 8-9 March 2018.
4. Ronald Poppe, "A survey on vision-based human action recognition", Image and Vision Computing, pp. 976–990, 2010.
5. A. A. Efros, A. C. Berg, G. Mori, and J. Malik, "Recognizing action at a distance", Proceedings International conference on computer vision ICCV, 2003.
6. H. Jhuang, T. Serre, L. Wolf, and T. Poggio, "A biologically inspired system for action recognition", Proceedings International conference on computer vision ICCV, 2007.
7. K. Schindler and L. van Gool, "Action Snippets: how many frames does human action recognition require?" in Proc. IEEE Conf. Computer Vision Pattern Recognition, Jun. 2008, pp. 1–8.
8. J. C. Niebles and L. Fei-Fei, "A hierarchical model of shape and appearance for human action classification", In Proceedings Computer Vision Pattern Recognition CVPR, 2007.
9. I. Laptev and T. Lindeberg, "Local descriptors for spatiotemporal recognition", Proceedings International conference on computer vision ICCV, 2003.
10. S. Ali, A. Basharat, and M. Shah, "Chaotic invariants for human action recognition", Proceedings International conference on computer vision ICCV, 2007.
11. R. Gilad-Bachrach, A. Navot, and N. Tishby, "Margin based feature selection-theory and algorithms," in Proceedings of the twenty-first international conference on Machine learning, pp. 43–50, 2004.
12. Y. Sun, S. Todorovic, and S. Goodison, "Local learning based feature selection for high dimensional data analysis," IEEE transactions on pattern analysis and machine intelligence, vol. 99, 2009.
13. B. Chen, H. Liu, J. Chai, and Z. Bao, "Large margin feature weighting method via linear programming," IEEE Transactions on Knowledge and Data Engineering, vol. 21, no. 10, pp. 1475–1488, 2009.
14. Dishani Lahiri, Chhavi Dhiman, "Abnormal Human Action Recognition using Average Energy Images", Conference on Information and Communication Technology (CICT'2017).
15. P. A. Dhulekar, Dr. S. T. Gandhe "Action recognition based on histogram of oriented gradients and spatio-temporal interest points", International Journal of Engineering & Technology (IJET), 7 (4), 2153-2160, 2018.
16. N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," In proceedings Computer Vision Pattern Recognition CVPR, pp: 886–893, 2005.
17. G. Varol, I. Laptev and C. Schmid, "Long-term Temporal Convolutions for Action Recognition", IEEE Transaction on Pattern Analysis and Machine Intelligence, Volume: 40, Issue: 6, June 1 2018 Pages. 1510 – 1517.
18. Wei Yang,Kuanquan Wang and Wangmeng Zuo, "Neighborhood Component Feature Selection for High-Dimensional Data", Journal Of Computers, Vol. 7, No. 1, January 2012.

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

AUTHORS PROFILE

Dr. Sanjay T. Gandhe is Principal of Sandip Institute of Technology & Research Centre, Nashik. He completed B.E., M.E. and Ph.D. (VNIT, Nagpur) in 1990, 1996 and 2008 respectively. He has published 27 research papers in International Journals and presented 41 papers in National/International Conferences. He is fellow member of IETE, IEI, Senior Member of IEEE and life member of ISTE. He has fetched several Research Grants from Governmental Apex Bodies. He is honored as Best Principal Award and Best Teachers in Professional Educational Category. He was Reviewer of IEEE Transaction on Dependable and Secure Computing in 2011-12. He served as Session Chair in various National and International Conferences. His area of expertise is Image Processing, Robotics and Machine Vision.

Pravin A. Dhulekar completed his bachelor degree in electronics and telecommunication engineering in the year 2007, M.Tech in 2011 and currently pursuing Ph.D from Savitribai Phule Pune University, Pune. He is research scholar of E&TC Dept. of Sandip Institute of Technology & Research Centre, Nashik. He has published 18 papers in International Journals and presented 24 papers in National/International Conferences. He is recipient of best paper award. He also served as reviewer for several reputed international journals and conferences. His research interests are Computer Vision and Machine Learning.