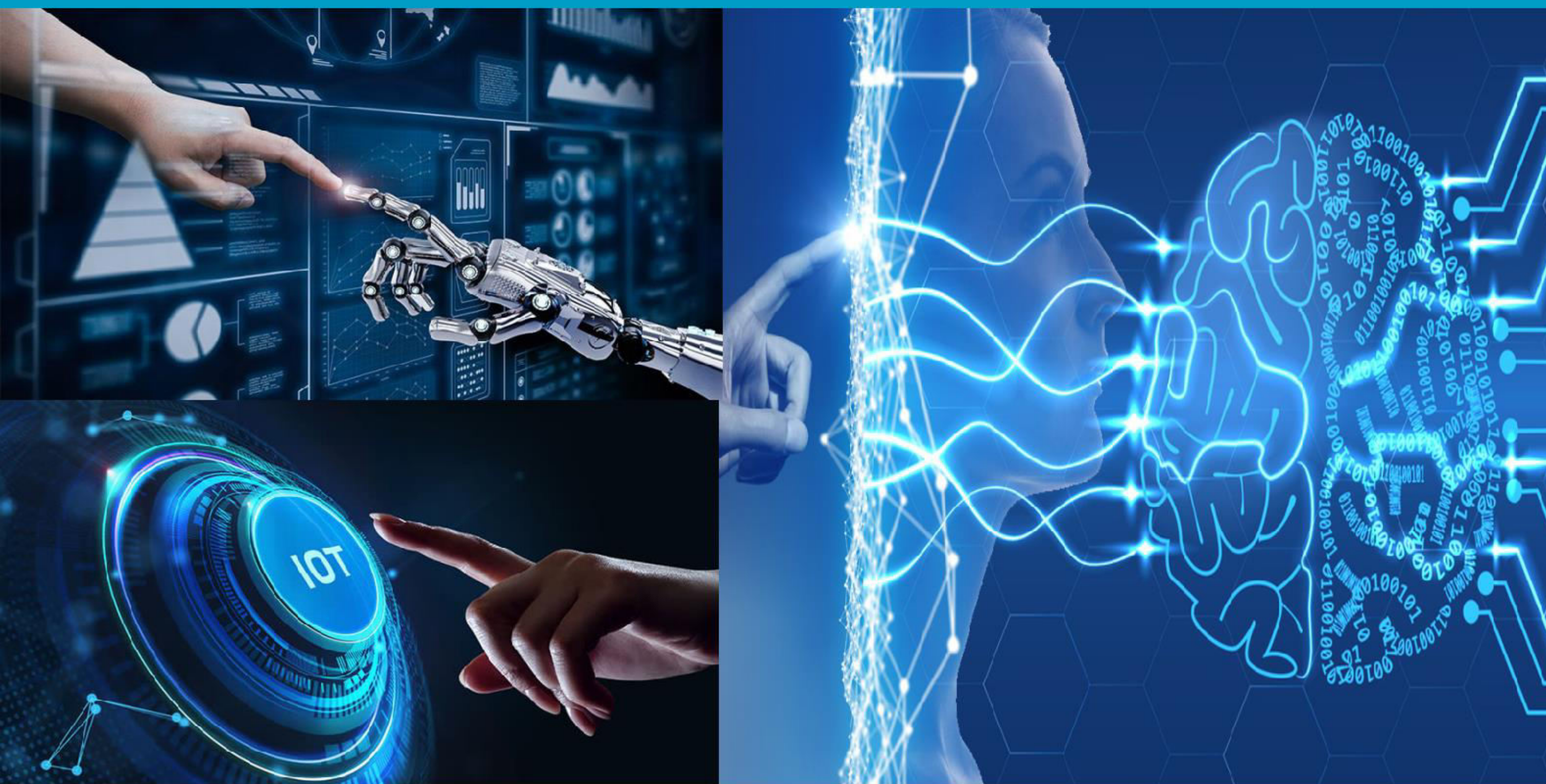
IJMRSETM

e-ISSN: 2395 - 7639



INTERNATIONAL JOURNAL OF MULTIDISCIPLINARY RESEARCH

IN SCIENCE, ENGINEERING, TECHNOLOGY AND MANAGEMENT

Volume 9, Issue 7, July 2022

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 7.580

+91 99405 72462 +9163819 07438 ijmrsetm@gmail.com www.ijmrsetm.com

| ISSN: 2395-7639 | www.ijmrsetm.com | Impact Factor: 7.580|

| Volume 9, Issue 7, July 2022 |

| DOI: 10.15680/IJMRSETM.2022.0907016 |

Spatio-Temporal and Support Vector Machine Based Human Action Detection

Monisa Nazir[1], Kirti Bhatia[2], Shalini Bhadola[3], Rohini Sharma[4]

[1]Student, Sat Kabir Institute of Technology and Management, Haryana, India

[2]Assistant Professor, Sat Kabir Institute of Technology and Management, Haryana, India

[3]Assistant Professor, Sat Kabir Institute of Technology and Management, Haryana, India

[4]Assistant Professor and corresponding Author, GPGCW, Rohtak, India

ABSTRACT: We have developed a model which can identify Human activities such as running, prodding, walking, hand clapping, boxing, and hand weaving. We have used spatial-temporal local features to extract the actions. Then we used a support vector machine (SVM) classifier to categorize different actions of a human. Our proposed method efficiently recognizes all main activities of humans.

KEYWORDS: Human Activity Recognition, Computer Vision, Spatial, Temporal, SVM

I. INTRODUCTION

Determining people's actions from videos or images is a complex procedure because of issues like contextual clutter, partial obstruction, differences in scale, outlook, illumination, and look. A multiple activity identification model is necessary for numerous applications, such as video monitoring systems, HCI, and robot for characterising human behaviour. To distinguish human behaviours, we have created a temporal and space local feature mechanism with a support vector classifier (SVM) machine learning approach. We used the KHH dataset and our method, and the outcomes are completely accurate.Examining actions from still photos or video clips is the aim of human activity recognition. This fact serves as the driving force behind person activity identification systems' quest to accurately classify input data into the relevant activity category. Human activities are divided into the following categories based on their complexity: gestures; (ii) single activities; (iii) inter-actions between humans or between humans and objects; (iv) collective actions; (v) behaviours; and (vi) events.

II. RELATED WORK

Human activity recognition is a part of a vision-based method like the identification of humans through a webcam [1-2], image quality enhancement [3], and recognition of face [4].Active learning (AR) in HAR is a recent research area. The AL methods seek to reduce the difficulty and expense of learning. A good quantity of revealing unmarked information samples should be chosen, and the labels should be requested. It increases prediction accuracy while reducing labeling effort [5].In contrast to other domains where active learning has gained popularity, only a small number of scholars have studied active learning in AR. To identify the most informative unlabeled data sample, Authors in [6] studied three distinct strategies and demonstrated that the annotation effort was decreased by a factor of 2.Reviews of techniques for recognising hand gestures using vision were published in [7]. These studies served to present the gesture identification as a means of communicating with devices and to give academics a rundown of developments in hand gesture identification in order to assist them pinpoint fields that still require more study. The authors of [8] concentrate on hand gesture identification technologies and also gesture taxonomies, presentations, and algorithms for recognising them. As opposed to actions and gestures, some academics are more concerned with interpreting behaviour and events.Motion estimation to identify human activities piqued the interest of other academics, and numerous reviews were published in this area. The authors of [9] make an effort to enumerate depth-based person

| ISSN: 2395-7639 | www.ijmrsetm.com | Impact Factor: 7.580|

| Volume 9, Issue 7, July 2022 |

| DOI: 10.15680/IJMRSETM.2022.0907016 |

movement analysis techniques. Person body part-based movement study, moving person tracing, and image sequence-centered person identity are three of the 3 sub topics of human movement analysis covered by [10].

III. AUTOMATIC RECOGNITION OF HUMAN ACTIONS

For content-centered video indexing, intellectual supervising, HMI, and virtual reality, human activity identification in videos are crucial. A promising use of computer vision is the automatic identification of human actions in video. Application scenarios include human-computer interaction, intelligent video surveillance, and content-based video retrieval. Although many scholars have conducted a lengthy study in this area, it is still difficult to identify human actions in films because of variations in scales, rotations, viewpoint, brightness, and occlusion in addition to geometric variances across intra-class items or motions. Extraction of features, visual representation, cross fusion, and classification are the four processes that makeup one of the most widely used frameworks for HAR.The classification procedure for recognising human activity based on spatial and temporal data is shown in Fig. 1.

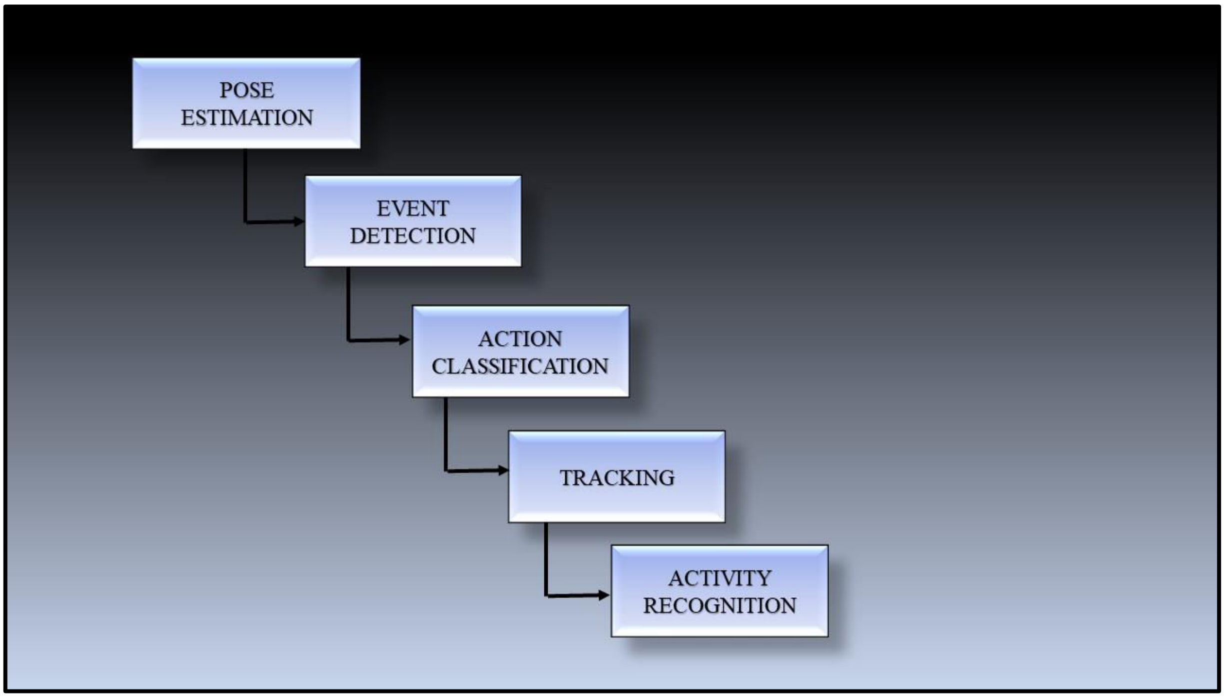


Fig 1: Steps of Human Activity Recognition

SPACE-TIME TRAJECTORY

Activity is seen as a collection of space-time trajectories in trajectory-based techniques. In these methods, a person's joints and body position are represented by 2-dimensional (XY) or 3-dimensional (XYZ) points. Depending on the action being performed, a person's joint positions alter when doing that action. Space-time trajectories are used to depict these changes as a 3D or 4D XYZT representation of the action. The body's joint location is tracked by the space-time trajectories in order to discern between various actions.

SPACE-TIME FEATURES

For the purpose of recognising human action, space-time features-based approaches extract features from space-time volumes or space-time trajectories. These traits are typically local in nature and contain action-specific discriminative properties. These properties can be categorised into two groups: sparse and dense, depending on the characteristics of space-time volumes and trajectories. While feature detectors based on optical flow are thought to be dense, those based

| ISSN: 2395-7639 | www.ijmrsetm.com | Impact Factor: 7.580|

| Volume 9, Issue 7, July 2022 |

| DOI: 10.15680/IJMRSETM.2022.0907016 |

on interest points are regarded as sparse. The majority of the newly proposed algorithms are built on these interest point detectors.

KTH HUMAN ACTION DATA COLLECTION

The Royal Institute of Technology in Sweden produced the KTH dataset [11] in 2004. This dataset includes 25 actors performing six various human activities in 4 separate contexts. Thus, there are 600 video sequences total, or 2564. This dataset is also regarded as being very basic for the evaluation of human activity detection algorithms because the films were taken with a static camera and background.

PROPOSED SOLUTION

With the help of local measurements in terms of spatiotemporal interest sites, we have shown in this work that action recognition is possible (local features). Such features can adjust to the magnitude, periodicity, and direction of movement sequences, capturing local motion events in videos and producing representations of video that are stable with regard to associated transformations.

Furthermore, local characteristics and SVM have been incorporated in a reliable classification method for spatial recognition. Corresponding to that, in this case, we investigate the integration of real time-space characteristics with SVM and use the subsequent methodology to recognise human actions. We provide a new video database for evaluation purposes and report findings from identifying six types of human activities carried out by 25 unique humans in various settings.

We build the scale-space representation of a picture sequence using a Gaussian convolution kernel to identify local features. The space and time variables of the related Gaussian kernel then determine the space and time neighborhood of features in space and time. By automatically choosing scaling parameters, the size of features can be adjusted to fit the Spatio-temporal extent of underlying image structures. Additionally, the features' shapes can be adjusted to match the local pattern's velocity, making them stable in the face of varying degrees of camera motion.The velocity and spatial representation of activities in consecutive frames are described by the spatiotemporal neighbourhoods of local characteristics. We compute spatio-temporal jets at the center of each feature using normalised derivatives to encode this data. Prior to computing descriptor l, we also warp the neighbours of features using predicted velocity values to ensure invariance with regard to related camera movements.A lexicon of basic events is produced by K-MCg of the descriptors l in the training set. A feature histogram is defined by the quantity of features with labels hi in a specific order (h_1, ..., h_n). When identifying motion in image sequences, we employ such histograms as one possible expression.

SUPPORT VECTOR MACHINES (SVM) BASED CLASSIFICATION

Modern big margin classifiers, such as the SVM, have lately grown in favor of the field of visual pattern recognition. Take into consideration the issue of categorising the training data set into two classes, where xi _N is a feature vector and yi 1, +1 its class label. The best hyperplane is the one that maximises the margin if we suppose that the two classes may be divided by a hyperplane of the form w x + b = 0 in some space H and that we are unaware of the distribution of the data. By employing Lagrange multipliers i(i = 1,... m), it is possible to find the ideal values for w and b by resolving a restricted optimization technique.

IV. EXPERIMENTATIONS AND RESULTS

The experimental setup, training procedure, and experimental outcomes for the suggested technique are covered in this section. On well-known action datasets like KTH, the suggested approach is tested.In relation to local characteristics, SVM classification has been integrated with motion descriptors (LF). Our concept was put into practice using

| ISSN: 2395-7639 | www.ijmrsetm.com | Impact Factor: 7.580|

| Volume 9, Issue 7, July 2022 |

| DOI: 10.15680/IJMRSETM.2022.0907016 |

MATLAB 2014 software. The KTH dataset's photos were then used to examine our results. We have selected a video frame from KTH database.Fig 2 shows recognition of walking activity.Fig 3 shows recognition of cycling by a human.Fig 4 shows the surfing activity of a human.Fig 5 shows recognition of boxing activity of humans.Fig 6 shows hand clapping by humans and Fig 7 shows hand waving by a human.Fig 8 and Fig 9 recognize jogging and running activities of human.

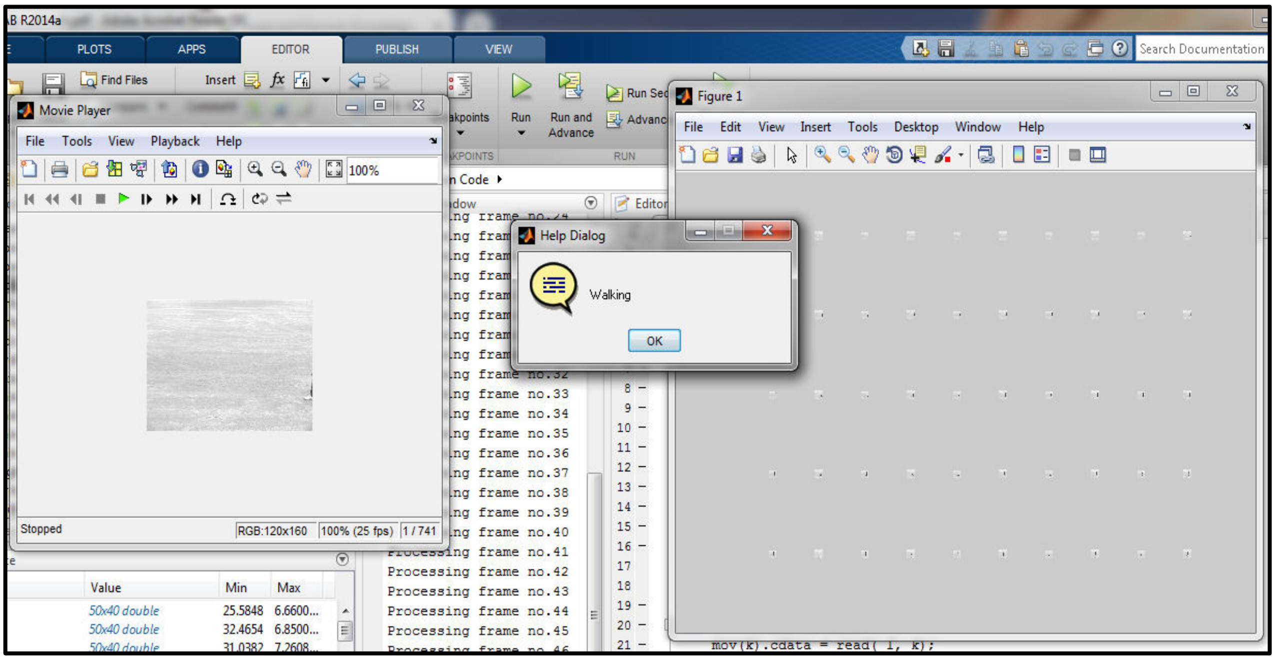


Fig 2: Recognition of Walking Activity

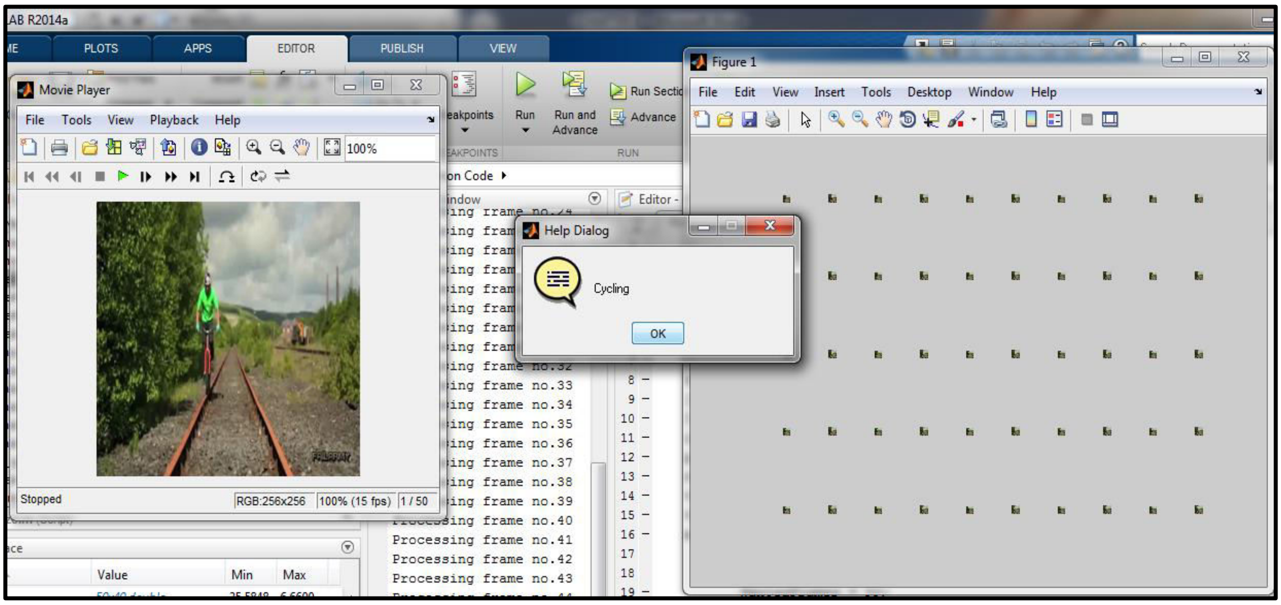


Fig 3: Recognition of Cycling Activity of Human

| ISSN: 2395-7639 | www.ijmrsetm.com | Impact Factor: 7.580|

| Volume 9, Issue 7, July 2022 |

| DOI: 10.15680/IJMRSETM.2022.0907016 |

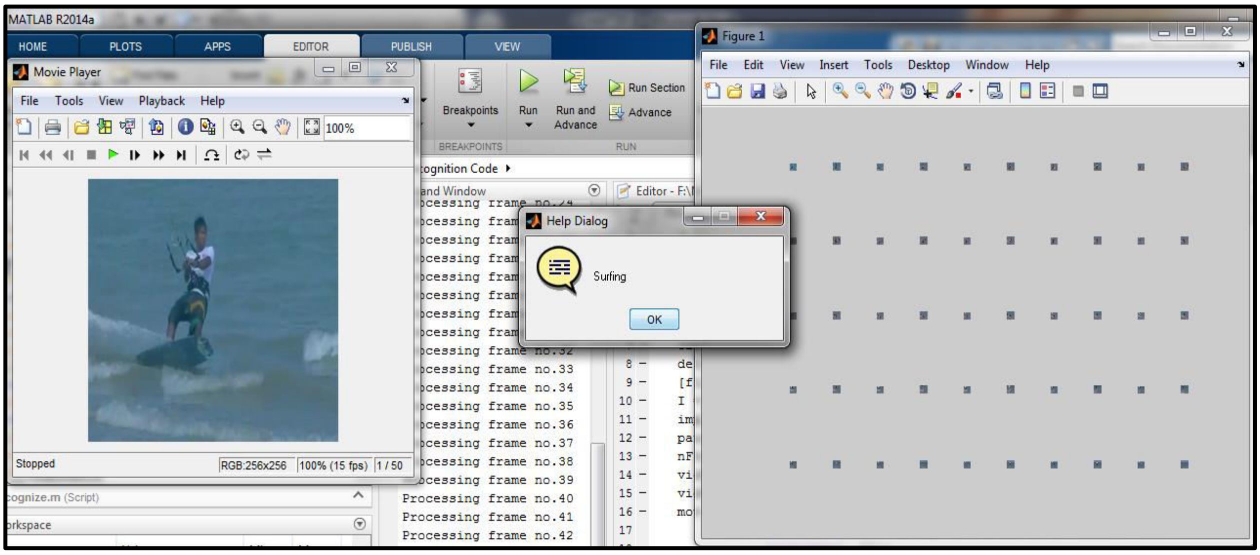


Fig 4: Recognition of Surfing Activity of a Human.

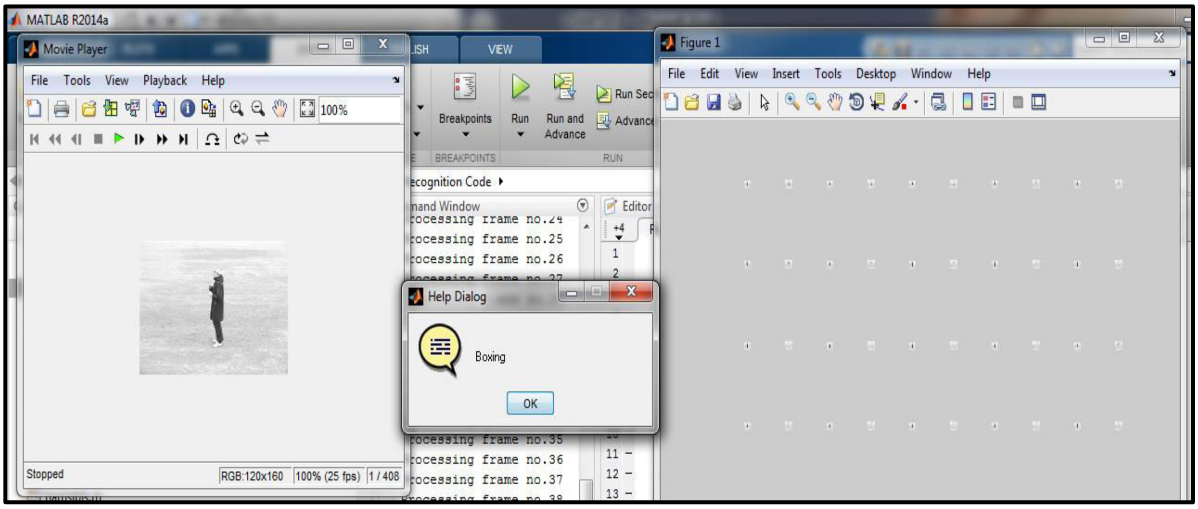


Fig 5: Recognition of Boxing Activity of Human.

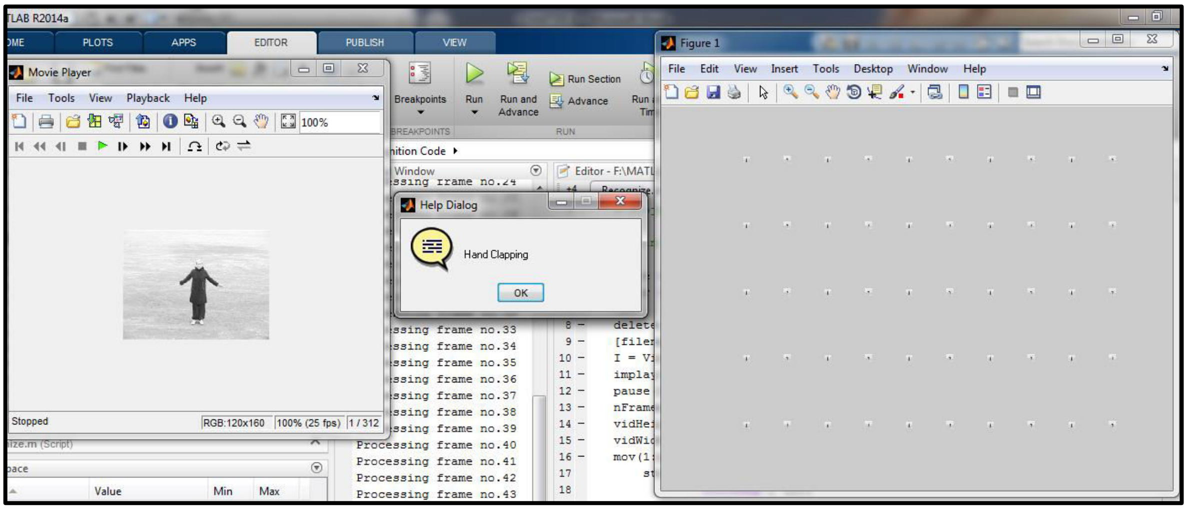


Fig 6: Recognition of Hand Clapping Activity

| ISSN: 2395-7639 | www.ijmrsetm.com | Impact Factor: 7.580|

| Volume 9, Issue 7, July 2022 |

| DOI: 10.15680/IJMRSETM.2022.0907016 |

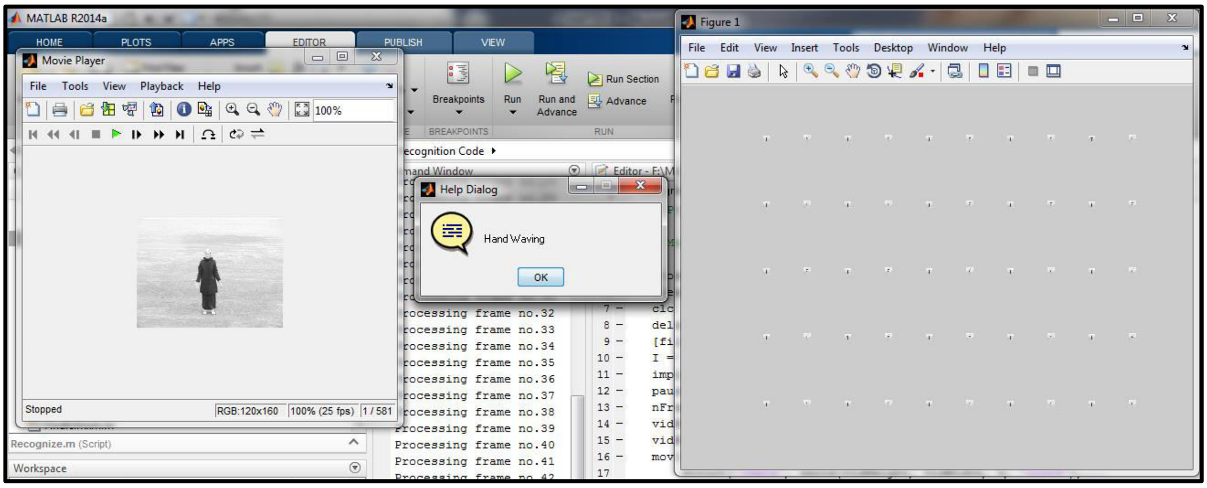


Fig 7: Recognition of Hand Weaving Activity

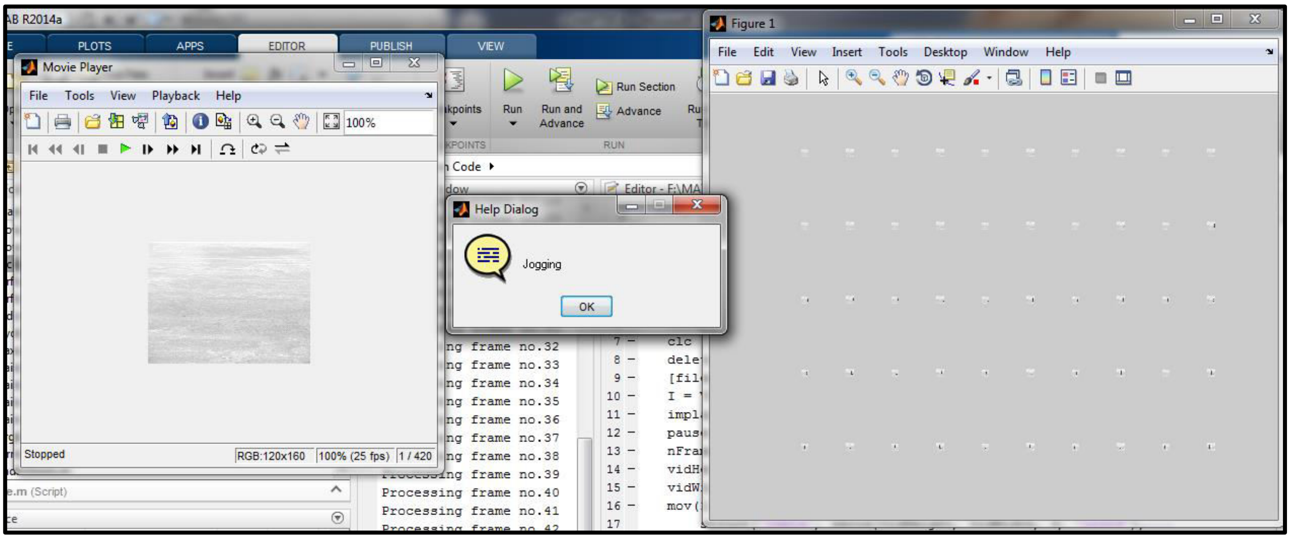


Fig 8: Recognition of Jogging Activity

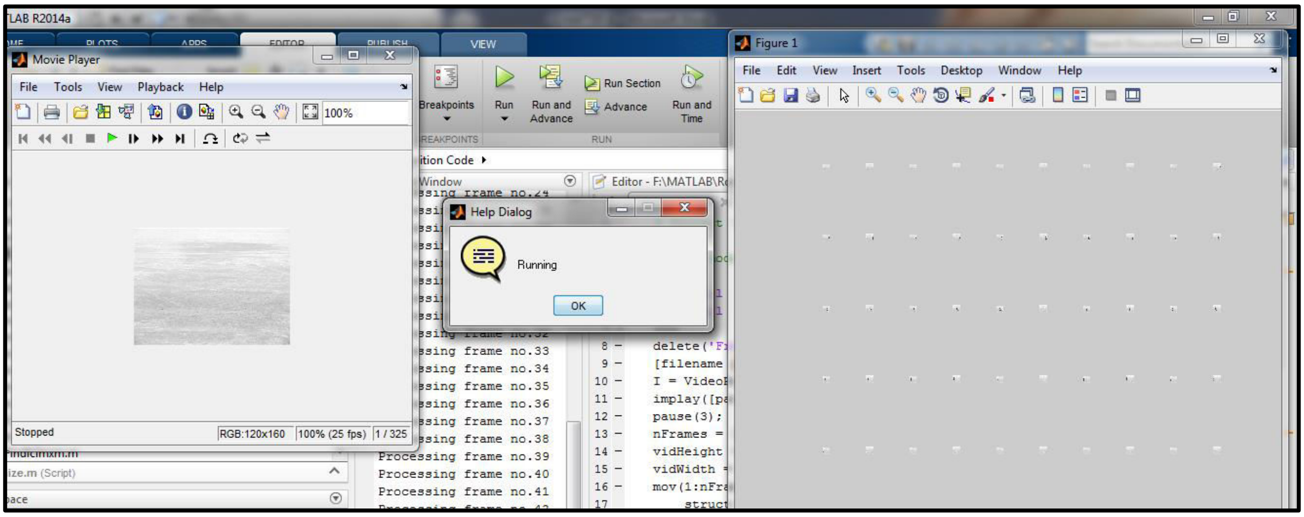


Fig 9: Recognition of Running

| ISSN: 2395-7639 | www.ijmrsetm.com | Impact Factor: 7.580|

| Volume 9, Issue 7, July 2022 |

| DOI: 10.15680/IJMRSETM.2022.0907016 |

V. CONCLUSION

Despite differences in clothes, lighting, and personal movement patterns, the localization of descriptors enables matching of similar events. But, some of the matched characteristics belong to various parts of (different) actions that are challenging to identify based just on local information because of the local nature of features and their corresponding jet characteristics. Therefore, it is clear that taking into account the spatial and temporal coherence of local characteristics will strengthen our strategy.We have shown the representation and recognition of motion patterns, such as human motions, using local spatiotemporal characteristics. We developed a novel motion recognition methodology that outperforms competing approaches in terms of recognition performance by merging local information with SVM.The benefits of representing gesture forms in regard of common attributs is that they can withstand changes in the pattern's scale, periodicity, and speed.

REFERENCES

[1] Nakul Nalwa, Shalini Bhadola, Kirti Bhatia, Rohini Sharma,Development of Vision Based People CountingFramework using Webcam, International Journal Of Multidisciplinary Research In Science, Engineering and Technology (IJMRSET), Volume 5, Issue 7, July 2022, pp. 1665-1671.

[2] Nakul Nalwa, Shalini Bhadola, Kirti Bhatia, Rohini Sharma,A Detailed Study on People TrackingMethodologies in Different Scenarios, International Journal of Innovative Research in Computer and Communication Engineering, Volume 10, Issue 6, June 2022, pp. 5907-5913.

[3] Mahesh Kumar Attri, Kirti Bhatia, Shalini Bhadola,Rohini Sharma, DEVELOPMENT OF A FOURIER TRANSFORMATION BASED MODELFOR IMAGE SMOOTHING, International Research Journal of Modernization in Engineering Technology and Science, Volume:04/Issue:06/June-2022, pp. 4204-4210.

[4] Sanju, Kirti Bhatia, Rohini Sharma, An analytical survey on face recognition systems, International Journal of Industrial Electronics and Electrical Engineering, Volume-6, Issue-3, Mar.-2018, pp. 61-68.

[5] Hossain, H. S., Khan, M. A. A. H. and Roy, N. Active learning enabled activity recognition. Pervasive and Mobile Computing, 2016.

[6] Alemdar, H., van Kasteren, T. and Ersoy, C.Active learning with uncertainty sampling for large scale activity recognition in smart homes. Journal of Ambient Intelligence and Smart Environments, 9, pp. 209–223, 2017.

[7] Cheng H, Yang L, Liu Z, Survey on 3d hand gesture recognition. IEEE Trans Circ Syst Video Techn2016, 26(9):1659.

[8] Rautaray SS, Agrawal A, Vision based hand gesture recognition for human computer interaction:A survey. ArtifIntell Rev, 2015, 43(1):1.

[9] Moeslund TB, Hilton A, Krüger V (2006) A survey of advances in vision-based human motion capture and analysis. Comput Vis Image Understand 104 (2-3):90.

[10] Kratz, L. and Nishino, K. (2012). Tracking pedestrians using local spatio-temporal motion patterns in extremely crowded scenes. IEEE transactions on pattern analysis and machine intelligence, 34 (5), pp. 987–1002.

[11] Liu,J., Yan,J., Tong,M., and Liu,Y.(2010). A Bayesian framework for 3D human motion tracking from monocular image, in IEEE International Conference on Acoustics, Speech and Signal Processing, Dallas, TX:IEEE, pp. 1398–1401.

INTERNATIONAL JOURNAL
OF MULTIDISCIPLINARY RESEARCH
IN SCIENCE, ENGINEERING, TECHNOLOGY AND MANAGEMENT

+91 99405 72462 +91 63819 07438 ijmrsetm@gmail.com

www.ijmrsetm.com