

Small Human Group Detection and Validation using Pyramidal Histogram of Oriented Gradients and Gray Level Run Length Method



Seemanthini K., Manjunath S. S.

Abstract: Over the decade's human detection in security and surveillance system became dynamic research part in computer vision. This concern is focused by wide functions in several areas such as smart surveillance, multiple human interface, human pose characterization, person counting and person identification etc. Video surveillance organism mainly deals with recognition plus classification of moving objects with respect to several actions like walking, talking and hand shaking etc. The specific processing stages of small human group detection and validation includes frame generation, segmentation using hierarchical clustering, To achieve accurate classification feature descriptors namely Multi-Scale Completed Local Binary Pattern (MS-CLBP) and Pyramidal Histogram Of Oriented Gradients (PHOG) are employed to extract the features efficiently, Recurrent Neural Network (RNN) classifier helps to classify the features into human and group in a crowd, To extract statistical features Gray Level Run Length Method (GLRLM) is incorporated which helps in group validation.

Keywords: Frame Generation, Hierarchical clustering, Multi-scale completed local binary pattern, Pyramidal histogram of oriented gradients, RNN and Gray level run length method.

I. INTRODUCTION

Human activity study using surveillance trajectory data is an interesting topic now a day. However action detection becomes a complicated problem because there will be variations in the action features, different action classes and jumbled background. The research has to be monitor for activity detection of a single individual and his interaction among other people in the group in reference to the above problems. Determining actions like walking, talking, hand shaking, etc are the well known events to be recognized. The work has been conducted in relation to the pedestrian detection in a group and persons validation in a small human group who are travelling together. In the literature there are several proposals are dedicated to modelling and understanding the human mobility with respect to other people reactions in the group by considering diverse aspects. Automatically identification and detection of multiple objects

is still remained an unsolved problem [01]. Multiple target interaction, Noise and occlusions present in each frame sequence leads to complexity in detecting object. So in order to decrease this complexity a novel methodology has been depicted here. The process of human crowd recognition and validating the human number in crowd involves the different processing steps such as individual frame generation, segmentation, extraction of features and classification of human and non human objects. Hence proposed model introduced to detect small human region and to count the human quantity in video sequence by classification. In this described model a proficient approach for small human group detection plus validation has been presented. A hierarchical based clustering algorithm is considered for the segmentation of the input frames. Multi-scale completed LBP (MS-CLBP) and PHOG descriptor used for the extraction of the features based on texture and shape features respectively with the intent to identify the human as well as non human objects [2][3]. To detect statistical features like detected human head quantity count, width between the detected individual heads, a proficient method (GLRLM) is used [4]. Proposed model validates the human group by using statistical features.

II. LITERATURE SURVEY

Weina Ge et.al [01] has discussed about the automatic human detection in crowd. Hierarchical clustering based on bottom-up approach has been depicted for the detection of the small individual group. The method uses the hausdorff distance which is defined in favour of velocity and pair-wise proximity. Here the model tells about the small pedestrian group recognition by using trajectories which are extracted by the dataset. Explanation is given in a superior way for the agglomerative clustering of human perception that is in the small pedestrian human group. Results are demonstrated for the automated tracking which is able to provide related characteristics of crowds. Detected trajectory information is enough to visualize arm gestures and gaze directions. Chen Chen et.al [02] has introduced a completed local binary pattern (CLBP) on scene classification of remote sensing land. In order to classify the leading texture features of multiple resolutions, improved multi-scale CLBP descriptor has been projected in this paper. Two separate executions of MS-CLBP have been done based on the kernel-based learning machine which is compared in accuracy, classification and computational complexity. Dataset is tested by using 19 class satellite images and 21 class land-use datasets.

Revised Manuscript Received on December 30, 2019.

* Correspondence Author

Seemanthini K*, Asst. Professor, Dept. of ISE, DSATM, Bangalore, India

Dr. Manjunath S. S., Professor & HOD, ATME, Mysore, India

© The Authors. Published by Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP). This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>)

Small Human Group Detection and Validation using Pyramidal histogram of oriented gradients and Gray level run length method

The model exhibits the increased performance. Yuanyuan Ding et.al [03] explained about the hierarchical and ensemble algorithms to detect object. PHOG descriptor extracts shift and scale invariant features of image. PHOG features are then fed to feed forward models ensemble classification network which reduces curse of feature dimensionality with improve object detection performance.

Hnin Mya Aye et.al [04] has told that the automated recognition of brain tumours in magnetic resonance images (MRI) is a tuff procedure owing to the variability and complexity of the location, size, shape and texture of these lesions. Because of the intensity similarities between brain lesions and normal tissues, some approaches make use of multi-spectral anatomical MRI scans and some other difficulties necessitate developing an approach that can detect tumour tissues using single spectral anatomical MRI images. The study of evaluating the efficiency of statistical feature over Gabor wavelet features using several classifiers has been discussed. This contribution fills the gap in the literature, as it is the first to compare these sets of features for tumour segmentation applications.

Alexandre Alahi et.al [05] proposed a quantitative research of identification based on crowded atmospheres like train stations, city centres etc. The model is implemented on standard dataset that contains Forty two million trajectories of train stations. The algorithms speaks about the problem of visualize pedestrian destinations and crowd mobility in large scale. Problems related to limited observations of sparse cameras and variations in pedestrian cues appearance across numerous cameras have been addressed here. In order to connect broken and unobserved trajectories, a latest descriptor Social affinity maps (SAM) has proposed in this paper. The tentative result shows the progress in the performance by using the SAM features and origin along with destination prior.

Chi-Chia Sun et.al [06] have been explained about the depth image information retrieval from stereo vision. Information about the depth image can be retrieved either by RGB-D camera or by dual camera, most of the times detection of objects mainly derived from the depth images or color information of the image. This paper combines the advantages plus disadvantages of both depth image and color information to achieve good result. The algorithm is capable of recognizing moving objects without background noise. The proposed model achieved 84.4% segmentation accuracy.

Marco Manfredi et.al [07] has described a way to analyse the crowded scenes. This paper is concentrating mainly on the human group gathering in same location for a while and static groups. SVM adopted for texture features classification. Crowd that contains spatial regions are filtered with intent to reduce the noise and false counting due to motion flows of human group. One inner texture descriptor and a SVM can be getting form single training set, which is further used for different models that analyse the crowd behaviour. Experiments results are conducted using general publicly available datasets and validated for the proposed system.

K.N.Tran et.al [08] has presented an efficient and novel framework for activity analysis of human group. Humans in a

group were recognized using undirected chart having vertices which represents people, edges among two people will be weighted on order to calculate how much they are interacting. To calculate degree of interface among people social-signalling cues are described here. A clustering technique based on graph detects interactive group. Two signalling cues are described and evaluated to discover group. BOW method is used to represent activity in group and SVM classifier for the action detection.

Pengming Feng et.al [09] have examined about the challenges like number variation, object occlusion and amount of noise, involved in human tracking. The model addresses the problems by Probability hypothesis density (PHD) filtering with respect to the implementation of Markov Chain Monte Carlo (MCMC). Social force model (SFM) an advanced approach is described to analyse the interaction among the targets. SFM is applied to validate the likelihood inside the MCMC sampling in favour of prediction step of PHD filter. One class support vector machine (OCSVM) is adopted to moderate the noise. The considered sequences are from the TUD, CAVIAR and PETS2009 datasets. The results are evaluated over the previous algorithms. The selected algorithms achieved good results.

Benoit Gauzere et.al [10] has proposed a new knowledge-based plus top-down approach for the tracking of human groups in the video sequence. First the model analyses the problems encountered in tracking and second it defines a top down approach rather bottom-up approach to predetermine the knowledge. The top-down approach allows us to keep consistency in trajectory extraction from the video sequence. The model diminishes the low level content of the less detected phases and it considered one frame per second rather than 7 from the original dataset. The preliminary validations are conducted on standard PETS 2009 dataset. The model states that combination of knowledge and low information are sufficient to track the human trajectories.

III. METHODOLOGY

Detection of human beings and validate the number of humans in the group is crucial task. The overall scheme of small human group detection and validation is illustrated in Fig. 1. The human zone is firstly predicted by frame by frame generation step. Recognition of human region includes computing of hierarchical based clustering followed by extracting features using Multi-scale completed local binary pattern (MS-CLBP) and Pyramidal histogram of oriented gradients (PHOG) descriptor methods. The next process is validating the quantity of humans by extracting their statistical features that calculates number of detected human heads and space between the individual head by using Gray level run length method (GLRLM). Each step is depicted in detail in the rest of the paper.

The proposed system is divided into 2 phases i.e. testing and training. In the same system setup the procedure is separated in two stages i.e. Rule1 and Rule2. Rule 1 gives the overall information about how to detect people. Rule2 explains about the group validation. The further sections describe Rule 1 and Rule 2 in detail.

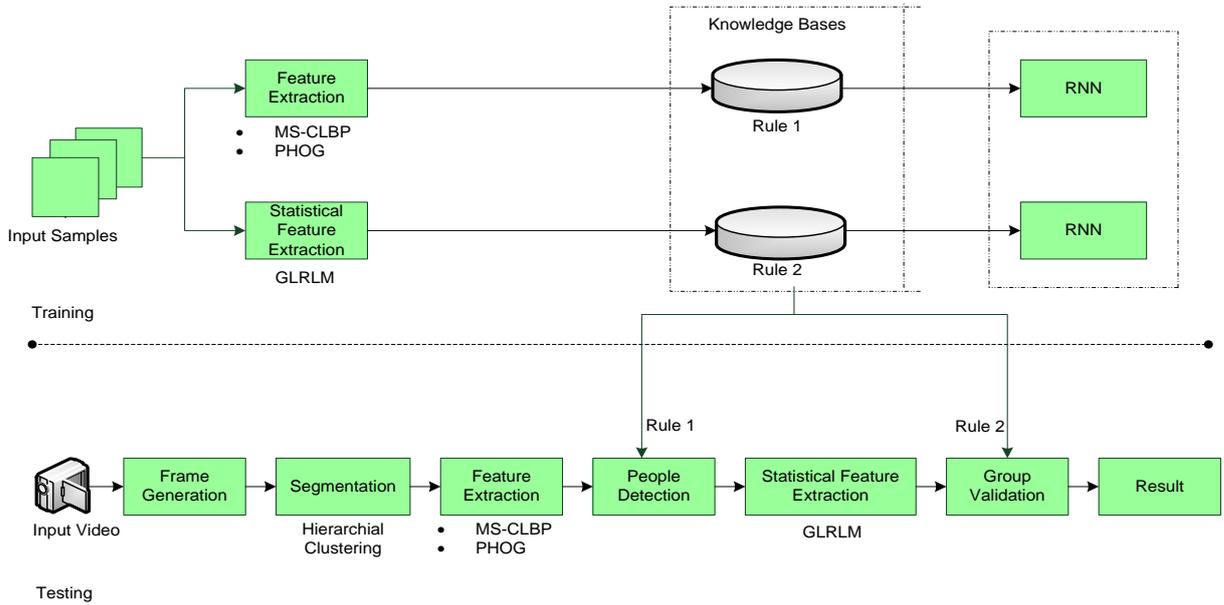


Fig. 1: Architecture of the Proposed System

3.1. Rule 1

First the video is splits into number of frames sequence by frame generation block. Each individual frames acts as input to the segmentation block. The hierarchical based clustering approach is considered for the further process. Hierarchical-clustering is employed because of its proficiency in crowd. Each clusters features will be extracted by using MS-CLBP and PHOG descriptor algorithms. These features will be compared with features of individual samples stored in the primary knowledge base. Human and non-human sections are separated by RNN.

3.1.1. Human Detection

Humans in the each frame are detected by employing hierarchical-clustering segmentation. The method begins with the each individual as separate clusters and merges two clusters to construct small human group gradually by strong inter-group proximity. The texture and shape features of each individual clusters are extracted by MS-CLBP and PHOG descriptor. Clusters are passed to RNN to differentiate human and non-human objects.

A. Segmentation

The brief review of this segmentation approach is discussed here. The hierarchical clustering process begins with the construction of a connectivity graph between the small human groups. Consider the small human group of size $k \geq 1$ and G_k is the connectivity graph corresponding to the group members. If the persons i, j are close for certain period of time, there will be an edge among the vertex η_i and η_j . i.e. $\rho_{ij} > T_t$. Here the method sets $T_t = 10$ for further experiments. The small human group helps to find the intra-group closeness by considering total edges in G_k as e_k and e_{k+1} as minimum edges which is desired in G_{k+1} after adding person p_i in G_k . p_i is included in existing group of size k if p_i is connected to the part of existing group members. i.e. the degree of $\eta_i \geq \lfloor \frac{k}{2} \rfloor$ have $e_{k+1} = e_k + \lfloor \frac{k}{2} \rfloor$. According to the definition $e_1 = e_1^{\wedge} = 0$. For $k \geq 1$, $e_2^{\wedge} = 1$ and $e_3^{\wedge} = 2$, Eq. (01) is derived,

$$e_k^{\wedge} = \begin{cases} \left(\frac{k}{2}\right)^2 & \text{if } k \text{ is even} \\ \frac{k-1}{2} \left(1 + \frac{k+1}{2}\right) & \text{if } k \text{ is odd} \end{cases} \quad (01)$$

The criteria for Intra group tightness matches by two small human groups G_p and G_q if it satisfies Eq. (02),

$$e_{p+q} \geq e_{p+k}^{\wedge} + (e_p - e_p^{\wedge} + e_q - e_q^{\wedge}) \quad (02)$$

The phrase in the parentheses shows extra edges of G_p and G_q contributing to intra-group tightness is merged group G_{p+q} . The larger the term in the parenthesis tells that small human groups are already tightly bounded with edges. In order to maintain a merged group to a tight group the requirement of edges in G_{p+q} is more.

B. Feature Extraction

The standard feature extraction algorithms MS-CLBP and pyramidal HOG (PHOG) descriptor are considered for the texture and shape features extraction of the input image which is described below.

a. Multi-Scale Completed Local Binary Pattern

MS-CLBP approach extract texture feature. Input image is represented into sign and magnitude. CLBP_M indicates the magnitude of the CLBP which is defined by Eq. (03) as follows,

$$CLBP_{m,r} = \sum_{i=0}^{m-1} p(|d_i|, c) 2^i, \quad p(u, c) = \begin{cases} 1 & u \geq c \\ 0 & u < c \end{cases} \quad (03)$$

c is the threshold value set to $|d_i|$ mean value, r indicates circular radius of neighbouring co-ordinate m . The pixel values will be in the decimal format after CLBP operation. CLBP_M (magnitude) and CLBP_S (signs) both could capture spatial information and texture such as corners and edges [2]. The algorithm fixes the amount of neighbour's m and tune varying r values for optimal combination achievement.

Small Human Group Detection and Validation using Pyramidal histogram of oriented gradients and Gray level run length method

Sign and Magnitude histogram features of CLBP are combined to form MSCLBP. In next iteration down-sampling of original image is done. Bi-cubic interpolation is used to obtain several images with the various scale difference.

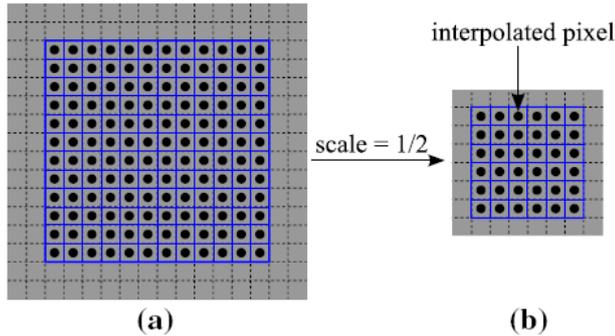


Fig. 2: (a) Original Image, (b) Down-Sampled Image

Down sampled image is shown in the Fig.2, which is half of original image, i.e. scale = 1/2. Finally CLBP_S and CLBP_M with constant radius and quantity of neighbours are applied to image of diverse scales. For each scale CLBP coded samples are generated and histogram features are computed from the two images. Lastly to form feature vector histograms are concatenated.

b. Pyramidal Histogram of Oriented Gradients

One more algorithm is employed to extract shape features from input samples, i.e. Pyramidal histogram of oriented gradients. To analyze object shape, edges of object are segmented. However edges are not good features to classify the object as edges varies with shape, rotation and illumination. In PHOG image splits into various blocks and at the output histograms are concatenated. Edge orientations are quantized with bandwidth and angles ranging from 0 – 180°. The extraction of PHOG based features are based on the following steps:

- Choose number of pyramid levels L (pyramidal length).
- For each pyramid level j given image is split into 4^j block therefore total block is calculated in Eq. (04)

$$\sum_{j=0}^L 4^j = \frac{4^{L+1} - 1}{3} \quad (04)$$

- Calculate the η -bin normalized HOG for each block.
- Calculate the pyramidal HOG to get the dimensional descriptor depicted in Eq. (05)

$$= \frac{\eta(4^{L+1} - 1)}{3} \quad (05)$$

The similarity measurement among the input pattern and prototypes at each pyramidal level is done based on the length of sub-descriptor at level j which is $\eta 4^j$, where η indicate bins number. Eq. (06), Eq. (07) and Eq. (08) measures the similarity.

$$y_{ij} = \sum_{k=1+\lfloor \eta 4^{j-1} \rfloor}^{\eta 4^j} \min(x_k, p_{ik}) \quad (06)$$

$$y_{ij} = \sqrt{\sum_{k=1+\lfloor \eta 4^{j-1} \rfloor}^{\eta 4^j} (x_k - y_{ik})^2} \quad (07)$$

$$y_{ij} = \sum_{k=1+\lfloor \eta 4^{j-1} \rfloor}^{\eta 4^j} x_k y_{ik} \quad (08)$$

Where y_{ij} defines similarity measure of input x with a prototype p_i at pyramid level j .

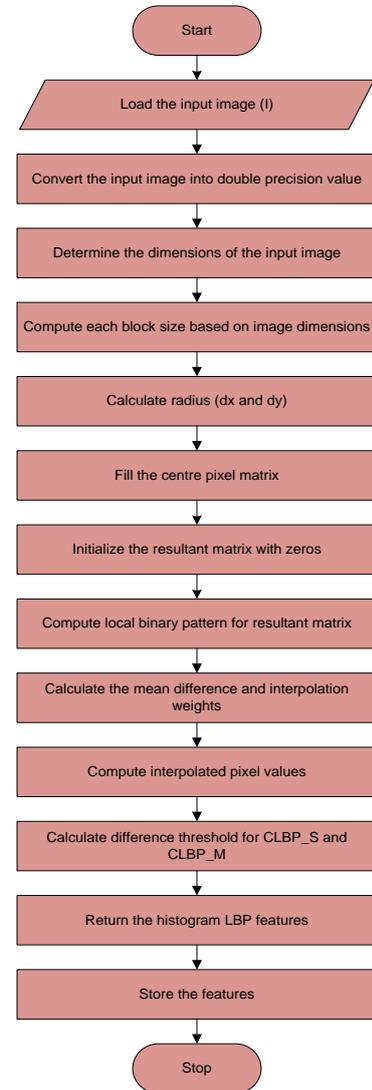


Fig. 3: Operational Flowchart of MS-CLBP

C. Recurrent Neural Networks

RNN are useful for sequence learning data, including NLP, speech and video respectively. The main intuition of RNN is to work on sequential information where the resultant output is based on the previous information. Generally in conventional feed forward neural networks input layers are independent to each other. RNN map history of previous inputs to each output. However their independency approach leads to insufficient prediction and classification in Natural Language Processing (NLP) because to predict the upcoming word there should be knowledge on previous word.

RNN make use of previous information to predict upcoming word so it is called 'recurrent'. RNN stores previous input in internal state memory. For a simple RNN, input sequence x of length is T is given in Eq. (09).

$$x = \{x_1, x_2 \dots \dots \dots x_T\} \quad (09)$$

Hidden layer h_t and output layer y_t can be calculated by

$$h_t = \sigma_h(w_h x_t + U_h h_{t-1} + b_h) \quad (10)$$

$$y_t = \sigma_y(w_y h_t + b_y) \quad (11)$$

w_h, w_y Denote the input-hidden and hidden-output weight matrices. U_h is the matrix between hidden layer and itself. b_h and b_y are biases. σ_h and σ_y are activation functions. Input is passed in Feed forward fashion at each time and then learning rule is applied.

Rule 1 concludes that the extracted-features from both MS-CLBP and Pyramidal HOG PHOG descriptor is compared with the already trained features in training phase based on which RNN differentiate the human as well as non-human objects.

3.2. Rule 2

The main goal of the second rule is to validate the quantity of humans in the detected group. The validation step is performed by employing a novel approach termed Gray level run length method (GLRLM) [4] in order to count number the human heads, extracting features from the human head pose and calculating the width between detected human heads. To validate the human group more accurately GLRLM represents salient object changes in the video frame. The more description about the algorithm is given below.

A. Gray level run length method

GLRLM extract statistical feature for analyzing texture. Texture information describes repetition of pixel intensity pattern. Run length compute number of adjacent pixel having similar intensity. GLRL Matrix is 2D matrix where each element is represented as $(p(i, j|\theta))$ j indicates number of elements with i as intensity. A 4×4 sub-image with 4 intensity levels is shown in Figure 4.

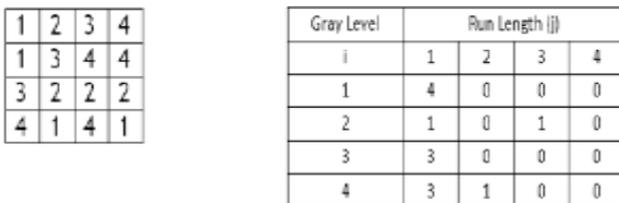


Figure 4: Matrix of image 4×4 pixel and GLRL Matrix

Algorithm : Pyramidal histogram of oriented gradients

Input: Input image either colour or gray, number of bins, angle and number of pyramid levels

Output: Pyramid histogram of oriented gradients

- Step.1: Convert the image from rgb to gray
- Step.2: For each block in the image
If $\text{sum}(\text{image}) > 100$
Calculate the gradient of the image
end
- Step.3: Calculate the square root of gradient of X and Y
- Step.4: Calculate the index of the gradient of X and Y
- Step.5: Determine both the angles of gradient X and Y
- Step.6: Calculate the bin matrix horizontal and vertical
- Step.7: In each step determine the number of zeroes in both vertically and horizontally
- Step.8: Calculate the phog descriptor

End algorithm

GLRL matrix is calculated by using $R(\theta) = [r'(i, l(\theta))]$. $r'(i, l(\theta))$ Compute how many runs are associated with image of length l . The GLRLM $R(\theta)$ is calculated for $0^\circ, 45^\circ, 90^\circ, 135^\circ$.

- SRE: Short run emphasis is given in the Eq. (10):

$$RF_1(R(\theta)) = \frac{1}{T_p} \sum_{i=0}^{G-1} \sum_{l=1}^{N_R} \frac{r'(i, l(\theta))}{l^2} \quad (10)$$

- LRE: Long run emphasis is given in the Eq. (11):

$$RF_2(R(\theta)) = \frac{1}{T_p} \sum_{i=0}^{G-1} \sum_{l=1}^{N_R} l^2 r'(i, l(\theta)) \quad (11)$$

- GLD: Gray level distribution is given in the Eq. (12):

$$RF_3(R(\theta)) = \frac{1}{T_p} \sum_{i=0}^{G-1} \left[\sum_{l=1}^{N_R} r'(i, l(\theta)) \right]^2 \quad (12)$$

- RLD: Run length distribution is given in the Eq. (13):

$$RF_4(R(\theta)) = \frac{1}{T_p} \sum_{i=0}^{N_R} \left[\sum_{l=1}^{G-1} r'(i, l(\theta)) \right]^2 \quad (13)$$

- RP: Run percentage is given in the Eq. (14):

$$RF_5(R(\theta)) = \frac{1}{T_p} \sum_{i=0}^{G-1} \sum_{l=1}^{N_R} r'(i, l(\theta)) \quad (14)$$

T_p is given by the Eq. (15).

$$T_p = \sum_{i=0}^{G-1} \sum_{l=1}^{N_R} r'(i, l(\theta)) \quad (15)$$

Thus, GLRLM depicts meaningful information to describe statistical features. Finally, these statistical features such as counting the number of human heads and calculating the width between the detected heads are compared and classified by RNN classifier with the previously stored features in knowledge base based on which the human validation is done.

IV. EXPERIMENTAL RESULTS

The above discussed methodology is implemented and demonstrated here. The available violence detection dataset is considered for demonstration. Algorithm has taken the video frames as the input generated from the video input. The video clip consists of 3 – 4 persons from the dataset. Finally proposed system proposed detects human and group. Figure 5 and 6 shows results. The frames are sized to 256×256 for processing each frame. The Figure 5(a), (b) and (c) shows the original frame, gray conversion and binary conversion frame respectively. To validate human Hierarchical clustering segmentation, PHOG and MS-CLBP features extraction algorithm is employed followed by RNN shown in Figure 5(d). To determine human group GLRLM feature extraction method is employed followed by RNN classifier as shown in Figure 5(e). Similarly the same procedure is considered for another video frames as depicted in Figure 6(a), (b), (c), (d) and (e) respectively. The proposed system achieves 91% of classification accuracy. Evaluation of system is done by computing True positive rate (TPR), Precision or Positive Predictive Value (PPV) and Accuracy using Eq. (16), (17) and (18) respectively.

Small Human Group Detection and Validation using Pyramidal histogram of oriented gradients and Gray level run length method

$$TPR = \frac{TP}{P} = \frac{TP}{TP + FN} \quad (16)$$

$$PPV = \frac{TP}{TP + FP} \quad (17)$$

Precision is nothing but mechanical or scientific exactness.

Accuracy is given by,

$$ACC = \frac{TP + TN}{P + N} = \frac{TP + TN}{TP + TN + FP + FN} \quad (18)$$

Accuracy is nothing but amount of exactness of a quantity or expression

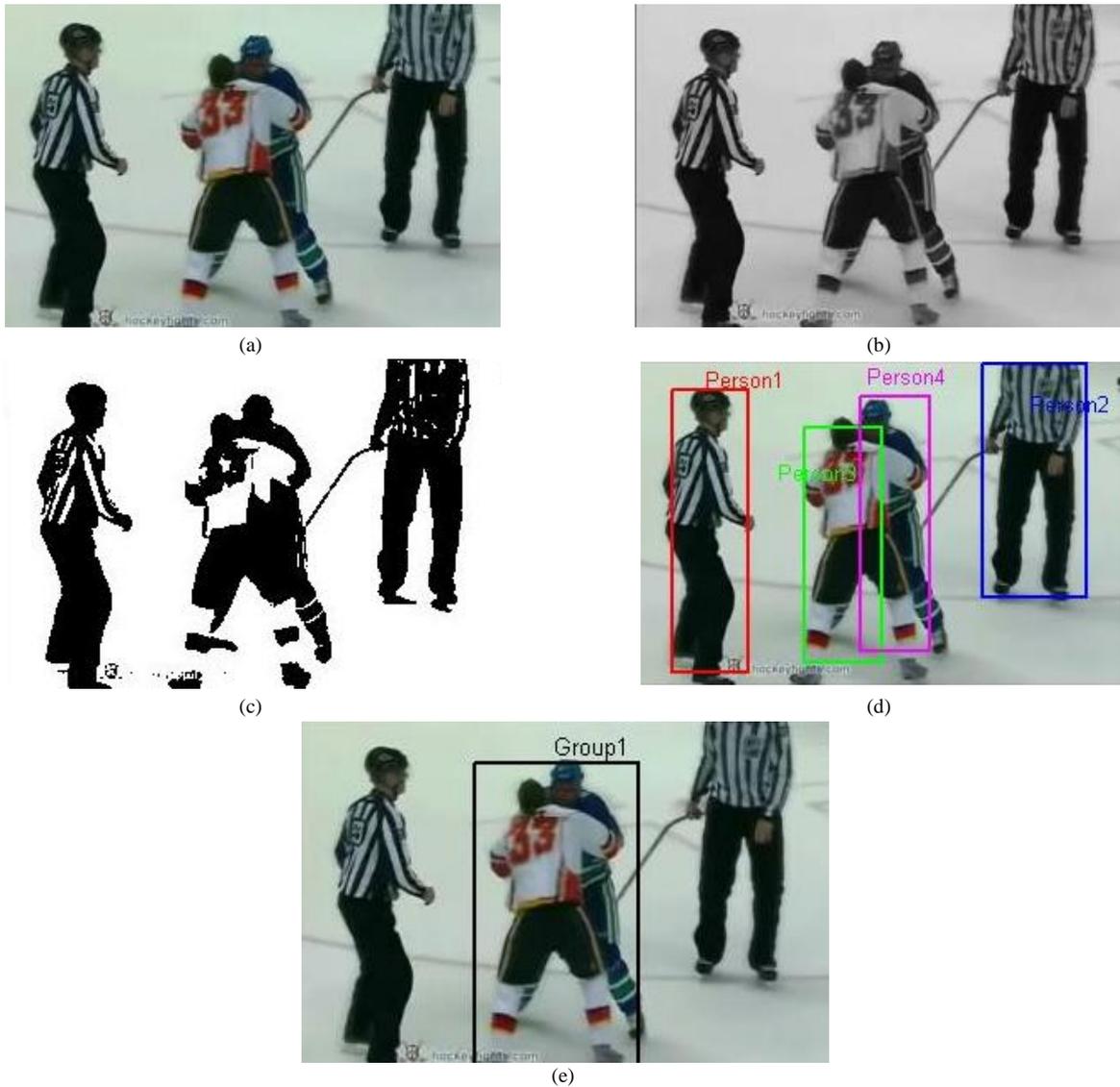
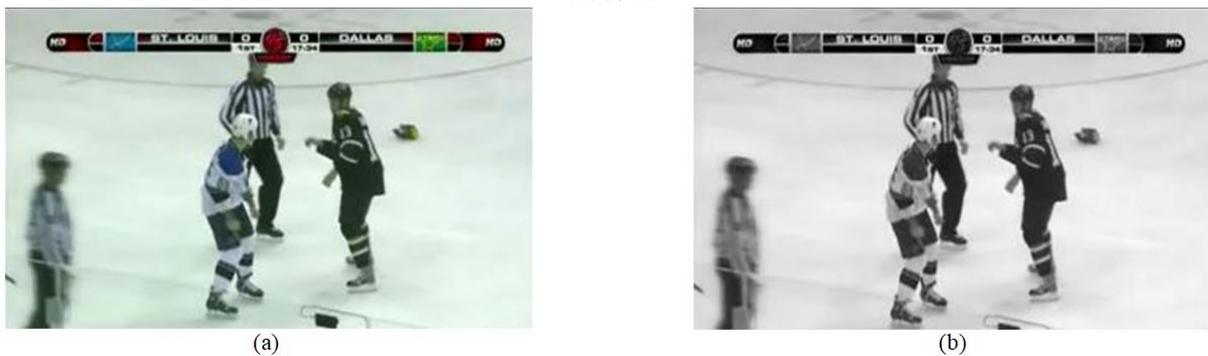


Figure 5: Resultant Output of Video Dataset 1; (a) Input Frames of Video 1; (b) Gray Conversion; (c) Binary Conversion; (d) Human Validation Results of Respective Frames; (e) Group Detection Result of Respective Frames of video 1.



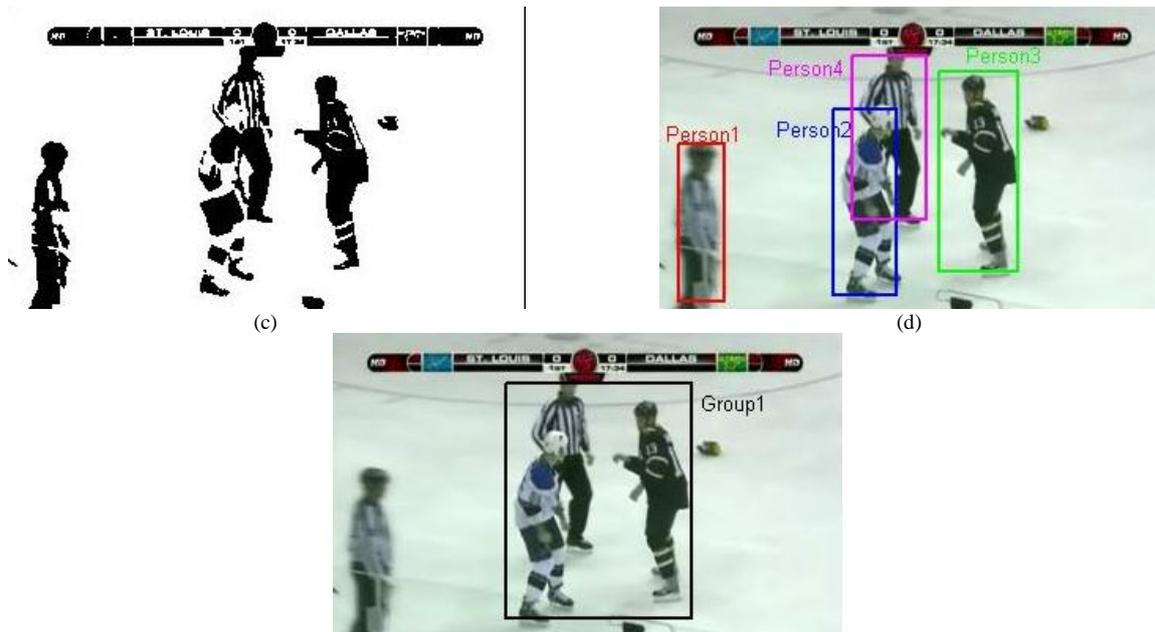


Figure 6: Resultant Output of Video Dataset 2; (a) Input Frames of Video 1; (b) Gray Conversion; (c) Binary Conversion; (d) Human Validation Results of Respective Frames; (e) Group Detection Result of Respective Frames of video 2.

V. CONCLUSION

We have represented a novel methodology for small human group detection and validation using pyramidal HOG (PHOG) and GLRLM. The Segmentation part has been achieved by using new algorithm termed hierarchical clustering. Segmented image features are extracted using MS-CLBP and PHOG descriptor. In order to validate the number of humans GLRLM statistical feature extraction technique is employed. The obtained features are then matched with the features stored in the knowledge base using RNN. This method is exposed to offer accurate small human group detection and validation.

REFERENCES

1. Ge W, Collins R.T and Ruback R.B, "Vision-Based Analysis of Small Groups in Pedestrian Crowds", IEEE, Vol. 34, No. 5, pp. 1003-1016, 2012.
2. Chen C, Zhang B, Su H, Li W and Wang L, "Land-use Scene Classification using Multi-Scale Completed Local Binary Patterns" Springer, Vol. 10, No. 4, pp. 745-752, 2016.
3. Murtza I, Abdullah D, Khan A, Arif M and Mirza S.M, "Cortex-Inspired Multilayer Hierarchy Based Object Detection System using PHOG Descriptors and Ensemble Classification", Springer, Vol. 33, No. 1, pp. 99-112, 2017.
4. Nabizadeh N and Kubat M, "Brain Tumors Detection and Segmentation in MR Images: Gabor Wavelet vs. Statistical Features", Elsevier, Vol. 45, pp. 286-301, 2015.
5. Alahi A, Ramanathan V and Fei-Fei L, "Socially-Aware Large-Scale Crowd Forecasting", IEEE, pp. 2203-2210, 2014.
6. Sun C.C, Wang Y and Sheu M.H, "Fast Motion Object Detection Algorithm Using Complementary Depth Image on an RGB-D Camera", IEEE, Vol. 17, No. 17, pp. 5728-5734, 2017.
7. Manfredi M, Vezzani R, Calderara S and Cucchiara R, "Detection of Static Groups and Crowds Gathered in Open Spaces by Texture Classification", Elsevier, Vol. 44, pp.39-48, 2014.
8. Tra K.N, Gala A, Kakadiaris I.A and Shah S.K, "Activity Analysis in Crowded Environments using Social Cues for Group Discovery and Human Interaction Modelling", Elsevier, Vol. 44, pp. 49-57, 2014.
9. Feng P, Wang W, Dlay S, Naqvi S.M and Chambers J, "Social Force Model-Based MCMC-OCSVM Particle PHD Filter for Multiple Human Tracking", IEEE, Vol. 19, No. 4, pp. 725-739, 2017.
10. Gaüzère B, Ritrovato P, Saggese A and Vento M, "Human Tracking using a Top-Down and Knowledge Based Approach", Springer, pp. 257-267, 2015.
11. Vázquez M, Steinfeld A and Hudson S.E, "Parallel Detection of Conversational Groups of Free-Standing People and Tracking of their Lower-Body Orientation", IEEE, pp. 3010-3017, 2015.
12. Nunes I, de Mel P.O.V and Loureiro A.A, "Group mobility: Detection, tracking and characterization", IEEE, pp. 1-6, 2016.
13. Mousavi H, Mohammad S, Perina A, Chellal R and Murino V, "Analyzing Tracklets for the Detection of Abnormal Crowd Behavior", IEEE, pp. 148-155, 2015.
14. Ali I and Daile M.N, "Multiple Human Tracking in High-Density Crowds", Elsevier, Vol. 30, No. 12, pp. 966-977, 2012.
15. Duan R, Fu and Kayaca E, "Tracking-Recommendation-Detection: A Novel Online Target Modeling for Visual Tracking", Elsevier, Vol. 64, pp. 128-139, 2017.
16. Ming-Xin Jiang, Chao Deng, Zhi-Geng Pan, Lan-Fang Wang and Xing Sun, "Multiobject Tracking In Videos Based on LSTM and Deep Reinforcement Learning", Complexity 2018.
17. Muhammad Irfan, Laurissa Tokarchuk, Lucio Marcenaro and Carlo Regazzoni, "Anomaly Detection in Crowds using Multi Sensory Information", IEEE International Conference On Advanced Video And Signal Based Surveillance (AVSS), pp. 1-6, 2018.
18. Sarvesh Vishwakarma and Anupam Agrawal, "A Survey on Activity Recognition and Behavior Understanding in Video Surveillance", Visual Computer, Springer, Vol. 29, No. 10, pp. 983-1009, 2013.
19. Yimeng Zhang, Weina Ge, Ming-Ching Chang and Xiaoming Liu, "Group context learning for event recognition", In 2012 IEEE Workshop on the Applications of Computer Vision (WACV), pp. 249-255, 2012.
20. Du Tran, Junsong Yuan and David Forsyth, "Video Event Detection: From Subvolume Localization to Spatiotemporal Path Search", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 36, No. 2, pp. 404-416, 2013.

Small Human Group Detection and Validation using Pyramidal histogram of oriented gradients and Gray level run length method

21. Tian Wang and Hichem Snoussi, "Detection of Abnormal Events via Optical Flow Feature Analysis", *Sensors*, Vol. 15, No. 4, pp. 7156-7171, 2015.
22. Yafeng Yin, Guang Yang and Hong Man, "Small Human Group Detection and Event Representation Based on Cognitive Semantics", *IEEE Seventh International Conference on Semantic Computing*, pp. 64-69, 2013.
23. Kyle Stephens and Adrian G. Bors, "Human Group Activity Recognition Based on Modelling Moving Regions Interdependencies", *In 2016 23rd International Conference on Pattern Recognition (ICPR)*, pp. 2115-2120, IEEE, 2016.
24. Naifan Zhuang, Tuorhongjiang Yusufu, Jun Ye and Kien A. Hua, "Group Activity Recognition with Differential Recurrent Convolutional Neural Networks", *12th IEEE International Conference On Automatic Face & Gesture Recognition (FG 2017)*, pp. 526-531, IEEE, 2017.
25. Kothapalli Vignesh, Gaurav Yadav and Amit Sethi, "A. Abnormal Event Detection On BMTT-PETS 2017 Surveillance Challenge", *In Proceedings of The IEEE Conference on Computer Vision and Pattern Recognition Workshops*, pp. 36-43. 2017.

AUTHORS PROFILE

Seemanthini.K received the B.E. degree in Computer Science & Engineering from People's Education Society College of Engineering, Mandya. and the M.Tech degree in Computer Science & Engineering from R.V.College of Engineering, Bangalore. She is currently pursuing research in Computer Vision and Pattern recognition under VTU, Belgaum. Her current research interests are anomaly event detection, Group Event detection and Action Recognition.

Manjunath S.S has received B.E degree in 2000 from Mysore University, Mysore and M.Tech degree in 2005 from VTU University, Belgaum, Karnataka, India. He completed Ph.D in Computer Science from University of Mysore, India. Currently he is working as a HOD & Professor at ATME,. His experience in teaching started from the year 2000. His areas of interests include microarray image processing, medical image segmentation and clustering algorithms.