THINK INDIA JOURNAL

ISSN: 0971-1260
Vol-22-Issue-3-July-2019

Fight Detection In Video Surveillance Using Kalman Filter & Blob Analysis

Pushpendra Kumar Pateriya
School of Computer Science & Engineering
Lovely Professional University, Phagwara, Punjab
Pushpendra.mnnit@gmail.com

AlishAhir
School of Computer Science & Engineering
Lovely Professional University, Phagwara, Punjab
Alisha.aheer@gmail.com

Dr.Parminder Singh
School of Computer Science & Engineering
Lovely Professional University, Phagwara, Punjab
Parminder.16479@lpu.co.in

Dr.Baljit Singh Saini
School of Computer Science & Engineering
Lovely Professional University, Phagwara, Punjab
Baljit.22078@lpu.co.in

Abstract— Human fight detection is a considerably important topic for video surveillance, and for strengthening automated security systems. Various methods have been used by researchers to detect the fight or violence in surveillance systems. Previously proposed methods do not support the efficiency-complexity trade-off. In In this paper, Kalman filter and Blob analysis are used to detect the fight of multiple persons in video surveillance. First of all, objects are created for reading the videos. Then, foreground detection and blob analysis are used to find out the important characteristics like area, bounding box etc. After that structure is used to keep the state of tracking the object. Kalman filter is used to detect and track the motion of objects. It takes the Euclidian Distance between the predicted center and center of finding. This method works well when the camera is stationary. This method has been applied to UCF crime dataset and significant accuracy has been achieved.

Index Terms—Kalman Filter, Blob Analysi,Figh Detection

I. INTRODUCTION

Nowadays, the use of surveillance CCTV cameras is very common in human life. We are using it for security and monitoring purposes.Incidences of violence and fighting are increasing day by day, so fight detection is one of the important areasof research. Automatic detection of human fighting in a surveillance video is a tedious task for surveillance.Various methods are being used by researchers to detect a fight or violence in video surveillance.Most of the researchers worked to detect two persons in fighting scenes. In this work, we detected the fight of multiple persons using Kalman filter and Blob analysis.Fight detection can be classified into two categories:gesture-based and frame sequence-based. For example feature point in the frame and position. Various methods are has been proposed, but the effectiveness achieved is less. Few of the methods are detecting a fight between two persons only. Few of them are detecting but time complexity is high. Such methods cannot

ISSN: 0971-1260
Vol-22-Issue-3-July-2019

be used in a real time environment. In this paper a novel method is proposed which is effective enough to detect human fighting.

The paper organization is as follows: Section IIcovers the literature review. Section III, contains the methodology with explanation. Section IV, shows the experimental results, and in Section V, the conclusion and future scope has been discussed.

II. LITERATURE REVIEW

As monitoring of violence in public places is becoming increasingly relevant to safety and security of common public, surveillance systems are now being deployed in public infrastructure and places such as educational institutions, offices and road sides etc.

Most of the existing surveillance systems are manually operated, and still need manual inspection and human intervention. The number of trained security workers isconsiderably less; commonlywe find a small group of security inspector who tracks tens to hundreds of video feeds with lesseffective alert system.

Humans are prone to disruptions and fatigue. After long periods of tracking the screens, security personnel can become less attentive.

Therefore, automated violence detection systems are increasingly needed today.In this paper an automated method is proposed to detect the fighting scene. Many of the researchers are working in the field of fight detection in video surveillance. Various methods are being used like motion co-occurrence, motion analysis, optical flow histogram, Hidden Markov Model (HMM), Gaussian Model of optical flow, Dynamic Bayesian Network (DBN), Oriented violent flows, AdaBoost Algorithm, Motion Weber local descriptor, Hough forest and convolution neural networks, motion analysis, Kalman filter, Morphological analysis etc.

The author in [1] used motion co-occurrence feature method to detect the fight in video surveillance. Block matching algorithm is used to extract the motion vectors. After that values of motion vectors are quantized separately. At last, direction and magnitude values are calculated based upon the existing and previous motion vectors. In this paper KNN is used for the classification purpose.

Author in [2] detected the fight based upon the motion analysis. First of all, Motion is detected and features are extracted. In this paper the fighting and non-fighting scenes are classified.

In paper[3], the author used the optical flow histogramto detect the abnormal events and violence. In this paper the log-polar histogram is used. PCA is used to decreases the dimensions of action representation. This method is applied on the video downloaded from the internet and works well on stationary camera.

In paper [4], the author used Hidden Markov model to detect a fight. Two models are used in this method, first one is ordinary model and second one is fight model.

ISSN: 0971-1260
Vol-22-Issue-3-July-2019

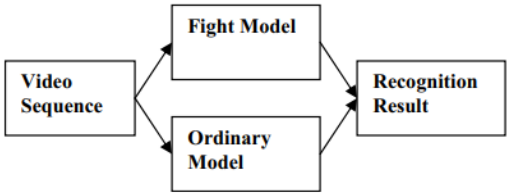

Figure 1 Frameworks for fight detection

In paper[5] author used Gaussian model of optical flow to extract the violent regions. The proposed orientation histogram of optical flow fed into Linear SVM which is used for the classification.

In paper [6], author presents a surveillance system CASSANDRA. The main goal of this system is to detect the human aggression in public areas. For this, audio and video signals both are combined in the fusion unit. Dynamic Bayesian Network (DBN) is used to estimate the level of aggression.

In paper [7], authorused oriented violent flows to detect the violence. Features are extracted using oriented violent flows. After that, different combinations are performed like feature extraction + classifier etc. AdaBoostalgorithm is also used to increase the performance.

Author in [8], proposed a method named as Motion Weber Local Descriptor. In this method, enhance the weber local descriptor by adding the non-permanent component to the appearance descriptor. To remove the redundant and irrelevant features kernel density estimation is employed on the motion weber local descriptor.

Author in [9], used Hough Forests and 2D convolutional neural network to detect the human fight in video. Features are extracted from the representative image frame which is taken from the video sequence. 2D CNN is used as classifier to classify the image and display the results.

In [10],author detected the fight in natural manner and low manner using motion analysis. First, noise removed from the original image. After that motion detected. In this paper, motion detection is classified into two categories: 1. Large Area 2. Small Area. A newmotion analysis technique is purposed which measure the strength of the various motion regions, which is known as motion region attraction.

In [11],the author is tracking objects using Kalman filter. In this paper, a comparison is also done between the Stauffer's adaptive background algorithm with spatio-temporal adaption of learning parameters and a Kalman-tracker in a feedback composition. In feed-forward, background module is giving the input to the Kalman-tracker. In backward, Kalman tracker adopts the learning parameters from the background module.

In [12], author usedKalman filter to track the object for underwater robot application. In this paper, Kalman filter is used to removing the problems of objects in SONAR images. The region of object is extracted using thresholding method and morphological operations.

In [13], the author used an extended Kalman filter to recognize and track the object for remote video surveillance.Recognition is done by analyzing by analytical approximation of the skeleton method extracted by the analyzed image. An extended Kalman filter is used for tracking.

ISSN: 0971-1260
Vol-22-Issue-3-July-2019

In [14], the author proposed a system to track the traffic system using morphological operations and blob analysis.The unnecessary things are reduced by converting the color into a grey color.Morphological operations are used to remove the small unnecessary objects. Blob analysis is used to extract the required information.

In [15], the author proposed a hybrid "learned/handcrafted" feature framework. the author used a deep learning paradigm to detect aggressive behaviour or human fight. Video sequence of input has been given to a 3D convolutional neural network that processes it.

By analysing the used methods we find out that to solve the problem Kalman filter and Blob analysis can be used. We have used them effectively in order to detect human fighting in a video effectively and timely.

III. PURPOSED METHOD

In this paper we have used Kalman filter for the navigation and aggregation because of its accurate estimation characteristics. Because of this characteristic, researchers manipulate its advantage into the tracking system. Now, it becomes the famous method to resolve the errors into the tracking system. It has two states. First one is correct and second one is predict. These states are used recursively. During the correct state, when correct steps are received then raw data is added into the filter. It predict state, it estimates the state and co-variance. Blob analysis is used to detect the 2D dimensional shape of the image. The detection is based upon some characteristics like centroid, area and bounding box. In this method, blob analysis is used to find the tracks which are connected in groups.

In the proposed method following steps has been followed in order to achieve the desired output:

1. Firstly, we have created the objects for analysis in the video, detecting the moving objects and showing the results.
2. Foreground detection is done and its output will be in binary mask. Pixel 0 values is used for background and pixel 1 value is used for foreground. Here, blob analysis is used to find out that tracks which are linked together by using the characteristics like area, centroid and bounding box.
3. Then, a group of tracks is created where every array will have an arrangement and moving the object. Structure is used where we can contain the state of track item. The defined structure is consists of 6 fields:
    - **Id:** it contains the digit id of the path.
    - **Bbox:**it is a bounding boxwhich contains the recent bounding rectangle of item used for displaying.
    - **Kalman Filter:** it is used for gesture based tracing. With the help of this, we track the motion of objects.
    - **Age:** it contains the information about the number of frames in which path was first noticed.
    - **Total visible count:** it shows the approximate frames of traced path.

ISSN: 0971-1260
Vol-22-Issue-3-July-2019

- **Consecutive invisible count:** it shows the total sequential frames for which the path was not noticed.

4. In this step, detection of objects is done. It takes the centroids and bounding box of the noticed object. Morphological operations are used to remove the noise and to fill the holes.
5. In this step, using Kalman filter we forecast the centroid of every path in the existing frame and inform its bounding container.First it predicts the existing position of the path and then shifts the bounding box so that its middle at the expected position.
6. In this step, we find out the Euclidian distance between predicted centre and centre of finding. It also comprises assurance of forecast which is sustained by Kalman filter.
7. In this step, it assigns the each path with equivalent recognition. Next, it stores the new bounding box and increase the age of path and full visible count by 1. Finally, the method sets the invisible count 0.
8. It symbols every unassigned path as invisible and enlarged by it age by 1.
9. Now, we delete the lost tracks. It also removed the newly produced paths that have been unseen for various frames overall.
10. In this step, we create the new tracks. Fresh tracks are created since non-allotted findings. The Kalman filter object is produced and also a new path is produced. Then, add it to the array of paths and increase the next id.

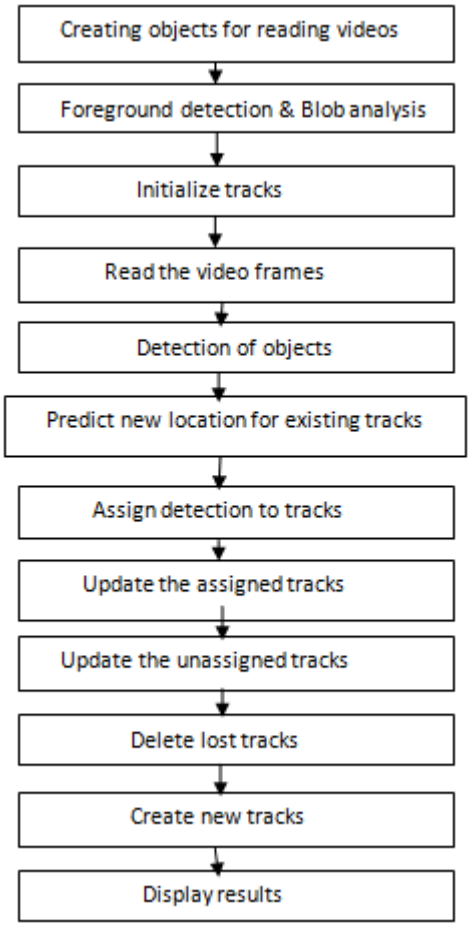

Figure 2 Block Diagram of proposed methodology

IV. EXPERIMENTAL RESULTS

ISSN: 0971-1260
Vol-22-Issue-3-July-2019

We have used UCF crime dataset and some downloaded videos from the internet. We have also taken the real-time simulated fighting videos from Lovely Professional University, Phagwara, Punjab. This method is applied on 100 videos. Input videos have been recorded in different environment. Time taking to detect a fight is dependent on the length of a video. We have also analysed that how much time it takes to detect the fight based upon the length of the video.With stationary camera we have achieved significant accuracy. This method achieved 90% accuracy at max. The minimum accuracy is 36% which is due to the fast movement in the position of camera. The graph between time duration and number of frames has been displayed in Figure 4. False positive and false negative ratio has been displayed in figure 5. The graph to depict the accuracy is displayed in Figure 6.

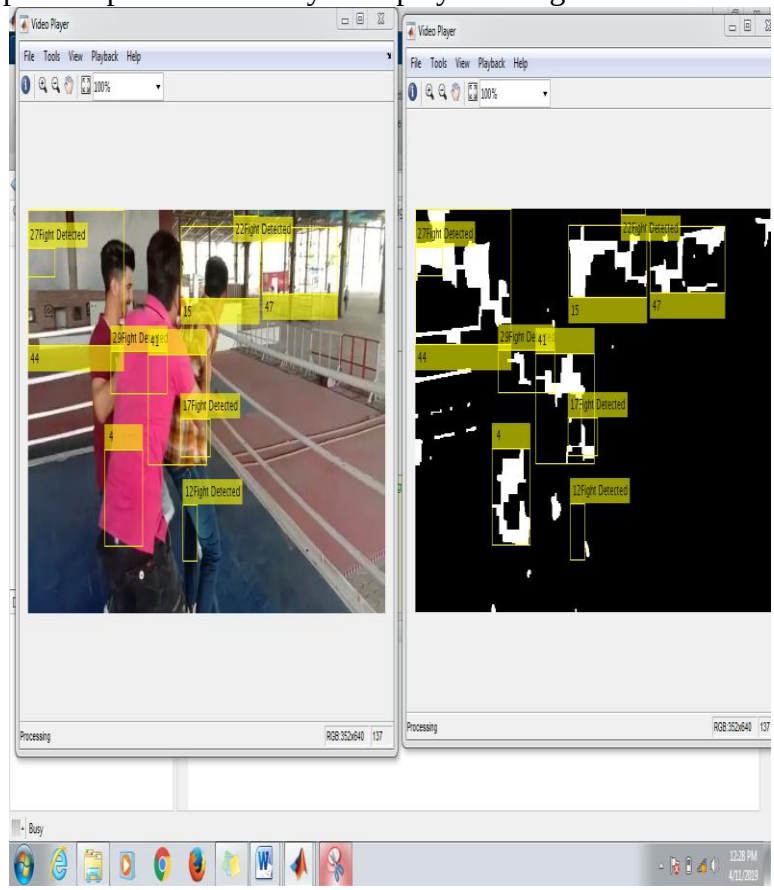

Figure 3: Fight Detection

ISSN: 0971-1260
Vol-22-Issue-3-July-2019

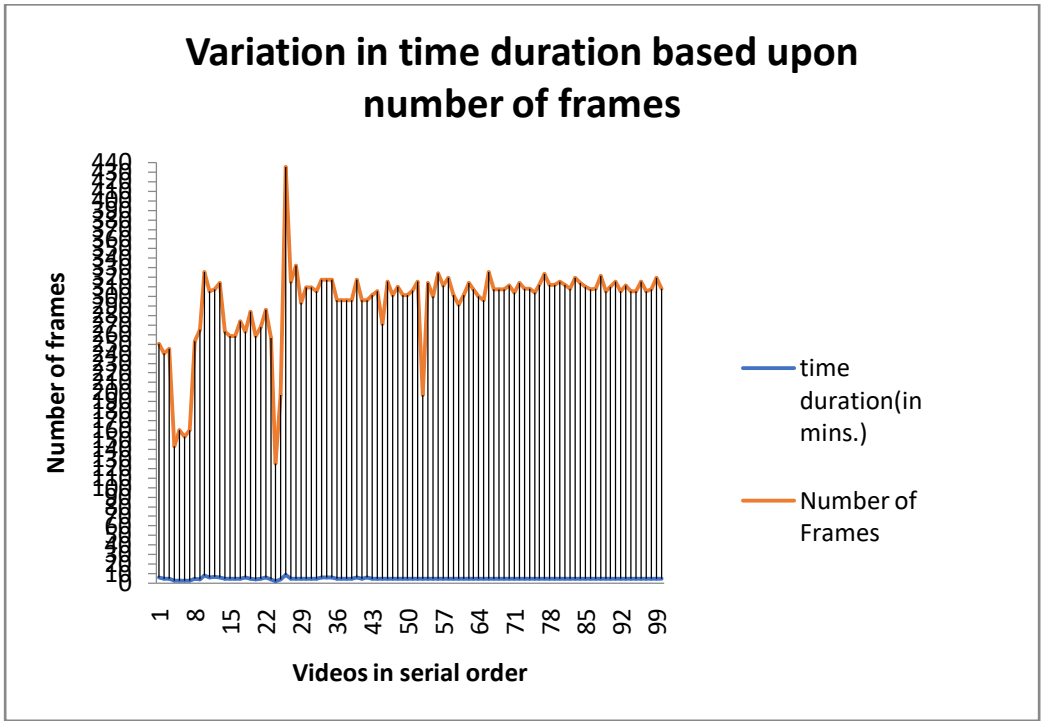

Figure 4: Variation in time duration based upon number of frames.

In fig.2, it shows that if number of frames is more than time duration will increase and if number of frames is less than time duration is decreases. So, we can say that time duration depend on the number of frames.
In fig.3, it shows the accuracy. This method achieved 90% accuracy at max. and minimum accuracy is 36% which is due to the fast movement in camera.

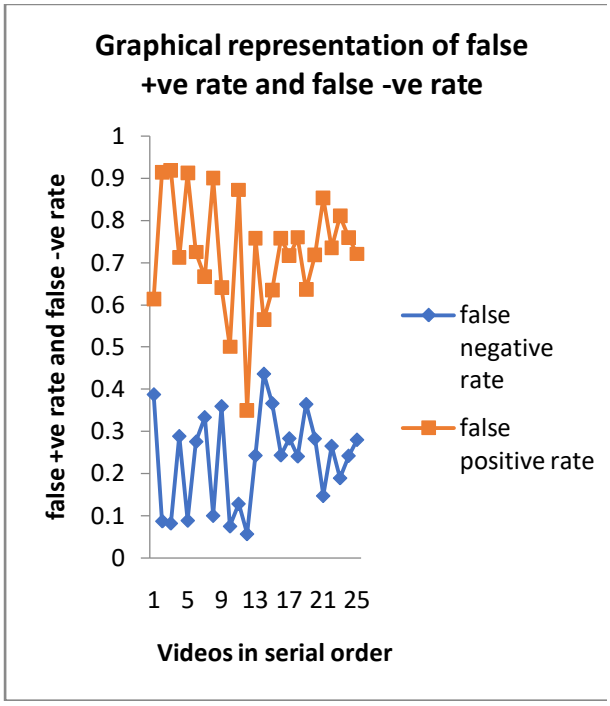

Figure 5: False positive rate and false negative rate

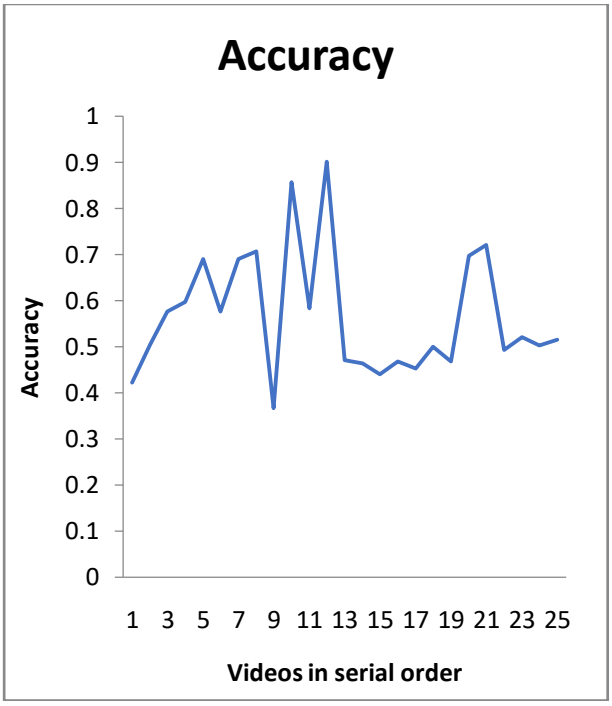

Figure 6: Accuracy

ISSN: 0971-1260
Vol-22-Issue-3-July-2019

Discussion: In this method, time duration is also depends upon the RAM and processor. If we use larger RAM and fast processor then time duration will definitely decrease. But if RAM is small and processor is slow then also it will work fine but take more time to detect the fight. We applied this method on 100 videos and some real - time videos. We also checked the time duration of this method like how much time it takes to detect the fight. These results are demonstrating that this approach works sensibly well in identifying the human fight scenes in videos. In previous research, only two persons were detected in the fighting scenes which are very limited. This method detects multiple persons during the fight in video surveillance. It also detects the two persons in the fight.

V. CONCLUSION

In this paper, an effective strategy for distinguishing fighting scenes from normal scenes in surveillance has been proposed. We used Kalman filter and Blob analysis to detect the fight in video surveillance. Some morphological operations are also used in this method. Kalman filter is used for tracking the fight activity scene.Blob analysis is used to find out the tracks which are linked together.As the meaning of viciousness is uncertain, in the most exceeding awful case setting, fight recognition is a challenging problem. In any case, there is still expectation in deciding an experimental reason for fight recognition. This approach can be connected straightforward to video reflection and can support high level video ordering in motion picture databases. This approach can work efficiently with larger RAM and fast modern processors. This method works well when camera is fixed or there is a minor movement in the camera. When there is a quick movement in the camera from one position to another then it does not detects the fight properly. So, in further we can enhance this work to remove this problem and develop a method to detect a sudden movement in a video, efficiently.

VI. REFERENCES

[1] E. Esen, M. A. Arabaci, and M. Soysal, “Fight detection in surveillance videos,” *Proc. - Int. Work. Content-Based Multimed. Index.*, no. June, pp. 131–135, 2013.

[2] E. Y. Fu, H. V. Leong, G. Ngai, and S. Chan, “Automatic Fight Detection Based on Motion Analysis,” *Proc. - 2015 IEEE Int. Symp. Multimedia, ISM 2015*, pp. 57–60, 2016.

[3] Y. Chen, L. Zhang, B. Lin, Y. Xu, and X. Ren, “Fighting detection based on optical flow context histogram,” *Proc. - 2011 2nd Int. Conf. Innov. Bio-Inspired Comput. Appl. IBICA 2011*, pp. 95–98, 2011.

[4] D. Liu, J. Wu, Y. Wang, J. Wang, and Z. Gong, “Fight detection based on hidden markov model,” *2012 IEEE 5th Int. Conf. Adv. Comput. Intell. ICACI 2012*, pp. 658–661, 2012.

[5] W. Jia, T. Zhang, B. Yang, J. Yang, X. He, and Z. Yang, “A new method for violence detection in surveillance scenes,” *Multimed. Tools Appl.*, vol. 75, no. 12, pp. 7327–7349, 2015.

[6] J. F. P. Kooij, M. C. Liem, J. D. Krijnders, T. C. Andringa, and D. M. Gavrila, “Multi-modal human aggression detection,” *Comput. Vis. Image Underst.*, vol. 144, pp. 106–120, 2016.

[7] Y. Gao, H. Liu, X. Sun, C. Wang, and Y. Liu, “Violence detection using Oriented VIolent Flows,” *Image Vis. Comput.*, vol. 48–49, pp. 37–41, 2016.

ISSN: 0971-1260
Vol-22-Issue-3-July-2019

[8] T. Zhang, W. Jia, B. Yang, J. Yang, X. He, and Z. Zheng, "MoWLD: a robust motion image descriptor for violence detection," *Multimed. Tools Appl.*, vol. 76, no. 1, pp. 1419–1438, 2017.

[9] I. Serrano, O. Deniz, J. L. Espinosa-Aranda, and G. Bueno, "Fight Recognition in Video Using Hough Forests and 2D Convolutional Neural Network," *IEEE Trans. Image Process.*, vol. 27, no. 10, pp. 4787–4797, 2018.

[10] E. Y. Fu and G. Ngai, "Automatic Fight Detection in Surveillance Videos."

[11] Nagamani Modalavalasa, G SasiBhushanaRao and K. Satya Prasad,"Target tracking system using kalman filter", International Journal of Computer Science, Engineering and and Information Technology IIJCSEIT), vol. 2, no. 2, April,2012.

[12] N. Modalavalasa, "An Efficient Implementation of Tracking Using Kalman Filter for Underwater Robot Application," *Int. J. Comput. Sci. Eng. Inf. Technol.*, vol. 2, no. 2, pp. 67–78, 2012.

[13] G. L. Foresti, "Object recognition and tracking for remote video surveillance," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 9, no. 7, pp. 1045–1062, 1999.

[14] Telagarapu, Prabhakar, MV Nageswara Rao, and Gulivindala Suresh. "A novel traffic-tracking system using morphological and Blob analysis." In *2012 International Conference on Computing, Communication and Applications*, pp. 1-4. IEEE, 2012.

[15] Serrano, Ismael, Oscar Deniz, Jose Luis Espinosa-Aranda, and Gloria Bueno. "Fight recognition in video using hough forests and 2D convolutional neural network." IEEE Transactions on Image Processing 27, no. 10 (2018): 4787-4797.