Multi Object Detection using Gaussian Mixture Model with Convoluted Moving Window Architecture Integrated by Kalman Filter for Background Subtraction Method

Ravindra Sangle

Research Scholar, PAHER, Udaipur, India

Ashok Kumar Jetawat

Professor, PAHER, Udaipur, India

Abstract:

Foreground object recognition in video is a critical step in several computer vision applications and automated video surveillance systems. Noise removal algorithms are used to detect mostly moving foreground objects. In recent decades, effective multi-object detection methodologies have been proposed for a variety of applications, like visual surveillance and behavior analysis. The range of applicants with the greatest performance and lowest cost is growing thanks to ongoing development in the technologies of video collection devices. Item detection methods are critical phases in objects recognition, surveillance, navigation systems. Object detection is the process of separating foreground and background items in photographs. Object tracking establishes the correlation among objects in a video sequence's succeeding frames. In this research, we offer algorithms that are split into 2 steps that includes Multi object detection utilizing Gaussian Mixture Model (GMM) and background suppression have performed. Multiple moving objects tracking using convoluted moving window Kalman filter (CMWKF). It can, however, cope with a variety of video sequences in the MOT 20 dataset. The experimental findings reveal that the proposed method detects and tracks foreground objects in complex and dynamic scenarios with high accuracy, robustness, and efficiency. This method also produces smoothened images without noise.

Keywords: multi-object detection, Gaussian Mixture Model (GMM), convoluted moving window Kalman filter (CMWKF), background suppression, video surveillance, video sequence.

1. Introduction:

Multiple Object Tracking intends to extricate directions of various moving articles in a picture grouping, is an essential errand in video understanding. A hearty and dependable MOT framework is the reason for a wide scope of uses including video observation, independent driving, and sports video examination [1]. Each following strategy requires an Object Detection instrument either in each casing or when the item initially shows up in the video. It handles division of moving articles from fixed foundation objects. This spotlights on more elevated level handling. It likewise diminishes calculation time. Because of ecological circumstances like light changes, shadow object division becomes troublesome and huge issue. A typical methodology for object identification is to involve data in a solitary edge. Notwithstanding, some item identification strategies utilize the worldly data figured from a grouping of edges to diminish the quantity of misleading recognitions. This worldly data is typically as casing differencing, which features districts that changes progressively in continuous casings [2].

GMM (Gaussian Mixture Model) is a limited blend likelihood appropriation model. The EM (Expectation-Maximum) calculation is an overall strategy for tracking down the Maximum Likelihood Estimation (MLE) from a deficient information and is really a superior drop calculation [3]. The benefit of GMM approach is that various kinds of burden dispersions can be genuinely addressed as an arched mix of a few ordinary appropriations with separate means and fluctuations. The issue of getting different blend parts (weight, mean, and difference) is formed as a Parametric Estimation issue. The EM calculation is an incredible asset in boundary assessment issues. It is an overall strategy for observing the Maximum-Likelihood gauge of the boundaries of a fundamental circulation from a given informational index when the information are inadequate or have missing qualities [4].

To manage various items following in unique scenes, a Kalman Filter Based Tracking calculation is utilized. The Kalman channel gives a powerful item following structure under questionable conditions. To deal with impediments, a few highlights are utilized. The covariance networks of interaction and estimation commotions are assessed consequently utilizing a moving-window assessment method. Specifically, process commotion covariance is tuned through a negative input conspire. Accordingly, underlying reaction and boundaries, as well as interaction and estimation commotion covariance, are all the while assessed in the recursive course of the channel [5]. The objective of this work is to follow the various items in the unique scenes. To accomplish this reason, this paper proposes the Kalman channel based numerous articles following calculation with tangled moving window. The commitment of this paper is as per the following

- To collect the various real time video surveillance data and convert the video into video frames.
- To detect multiple object in the video frames using Gaussian mixture model (GMM) which is integrated with convoluted moving window architecture and kalman filter (CMW_KF), smoothened images are then applied to background subtraction method with moving window which gives detected object present in background image.

The organization of the paper is as follows: In Segment 2, Works related to the proposed model is simply explained. The suggested model is described in depth in Segment 3. The experimental findings and some discussion are then reported in segment 4. Finally, segemnt 5 contains the paper's conclusion.

2. Related Works:

Moving article location and following utilizing Kalman channel was depicted in [6]. The following is vital for various item. The articles are followed the assistance of Kalman channel. This channel is utilized for the pixel astute deduction of current edge. As well as additionally used to be figure out the blunder between genuine place of the ball and assessed position esteem with the assistance of this channel. The item following technique on video transfers from genuine street climate was created and assessed by [7]. They contrasted the outcomes and other following calculations. This approach can follow the item with continuous impediment issues in a negligible measure of time in correlation with other existing calculations. A crossover model that further develops location precision with low computational time was introduced in [8]. The information picture is rearranged and bunched at different scales utilizing SLIC and K-implies grouping calculations separately. Gaussian Mixture Model (GMM) is created on different shading parts of the advanced picture at various scales. The vital highlights of this technique are object culmination and proficiency as far as computational time. Little notable highlights in the pictures are not distinguished. In [9] develops with mass age in view of highlights removed from submerged scenes utilizing the Bi-layered Empirical Mode Decomposition (BEMD) Algorithm. These separated elements are then familiar with the nonexclusive Gaussian Mixture Model (GMM) to give a premise to an item observing framework. This approach is contrasted and condition of-craftsmanship object identification conspires and dissected both subjectively and quantitatively and ended up being effective. An adjusted Gaussian Mixture Model (GMM) and Adaptive thresholding planned in [10]to further develop object recognition exactness for the open air reconnaissance. Inborn and extraneous upgrades in conventional GMM will deal with the open air dynamic scenes. The test results reason that this calculation productively distinguish objects in unique conditions as well as handle fractional impediments and certain measure of shadows proficiently. A Multi-Dimensional Kalman channel (MDKF) calculation is proposed in [11] for object following and movement identification. The mathematical examination of this following calculation accomplishes cutthroat following execution conversely, with best in class following calculations prepared on standard benchmarks. The state assessment with obscure non-fixed Heavy-Tailed Process And Measurement Noises (HPMN) is considered in [12]. A pointer is acquainted with recognize whether it is produced by the ostensible covariance or the bigger covariance by utilizing Kalman channel. Tries different things with manufactured information and genuine shrewd vehicle information show the viability of this channel under obscure non-fixed HPMN. The benchmark for quite a long time Tracking, MOT Challenge, was sent off in [13] with the objective to lay out a normalized assessment of Multiple Object Tracking Methods. The test centers around numerous individuals following, since walkers are very much examined in the following local area, and exact following and recognition has high down to earth significance. This allows to opportunity to assess best in class techniques for quite a long time following while dealing with very packed situations.
When the illumination changes there will be changes in the scene, and fails in the preservation of edge computing in moving object. Mistracking is occurred in the object’s appearance and disappearance because of the occlusion. These are the drawbacks in most of the existing system. To overcome these limitations, Gaussian mixture model (GMM) which is integrated with convoluted moving window architecture and kalman filter (CMW_KF) is proposed in this paper.

3. System Model:

This study provides a multi-object identification technique based on the Gaussian Mixture Model, as well as a convoluted movable windows Kalman filter approach for object tracking. In perplexing scenarios, this approach provides effective monitoring of many moving objects. Figure 1 represents the Gaussian mixture model with convoluted moving window architecture integrated by kalman filter. MOT20 is the real time video dataset taken in this model, this video is converted into small video frames. Multiple moving objects are detected by GMM model and noises will be present in the frames of the detected object. The noise is removed and smoothened by using Convoluted moving window kalman filter. After the removal noise the video frames are processed and analyzed.

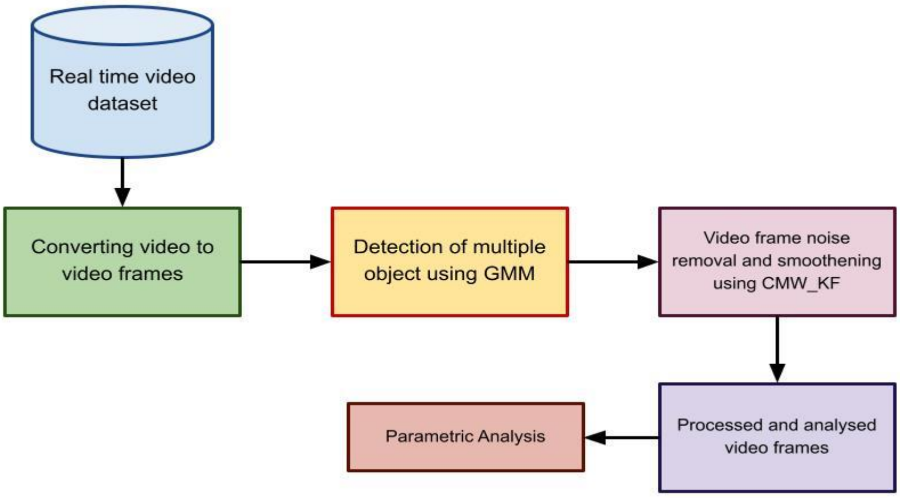

Figure 1: Gaussian mixture model with convoluted moving window architecture integrated by kalman filter

A parametric probability density function, GMM is a parametric model that estimates the probability distribution function based on a variety of object characteristics. When there is strong occlusion and dynamic scene changes, it is difficult to recognize several moving objects in computer vision technologies. Background modelling is always the initial stage in implementing background removal method to detect moving objects from each segment of video frames. The Gaussian Mixture Model may be used for this background modelling. To achieve the intended outcome, all incoming frames from a video series are subtracted from a reference background modelling frame and the difference is compared to a threshold value.

The general formula for Gaussian Mixture Model [14] is expressed in equation (1) as follows

$P(x_t) = \sum_{i=1}^{k}(\omega_{i,t} n(X_t; \mu_{i,t}, \sum i, t)$ (1)

Where, $\sum_{i=1}^{k}(\omega_{i,t}) = 1$

The mixture equal mean is given in equation (2) is

$\mu_t = \sum_{i=1}^{k} \omega_{i,t}\mu_{i,t}$ (2)

Background subtraction is a common technique for recognizing moving objects in films captured by static cameras. Assume that the backdrop has previously been updated for a certain sequence of frames before beginning the tracking procedure in the suggested technique. It is written as follows in equation (3):

$$B_t(x,y) = \begin{cases} 0 \; if \; |I_t(x,y) - b(x,y)| < \tau_B \\ 1 f \; |I_t(x,y) - b(x,y)| \geq \tau_B \end{cases} \quad (3)$$

Where $b(x,y)$ is the background that is already obtained. $I_t(x,y)$ is the gray image at time t. τ_B gives threshold value. Equation (1) and (2), are enough for detecting multiple moving object in videos. The$B_t(x,y)$) has information greater than $P(x_t)$relatively, as $P(x_t)$ has many noises caused by lightness.

The Kalman filter is a technique that estimates the parameters of a system and predicts future observations based on noisy measurements collected over time. It makes predictions, takes a measurement, and then updates itself depending on the forecast and measurement comparison at each time step. It is a mathematical evaluator that predicts and updates the states of a large number of linear processes.

Processing equation is expressed in equation (4) as follows

$x_k = Ax_{k-1} + w_{k-1}$ (4)

Where A is the transition matrix and x k denotes the state from time k-1 to time k. The Gaussian process noise N(.) vector w (k-1) has the following normal probability distribution p. (w). Consider a tracking system in which X_k is the state vector, which represents the object's location, velocity, and dynamic behaviour, and subscript k indicates discrete time. The basic goal is to use the measurement Z_k to estimate the value of X_k. Equation (5) represents the probability distribution function, whereas equation (6) reflects the observed measurement.

$p(w) \sim N(O, Q)$ (5)

$Z_k = Hx_k + V_k$ (6)

Here H denotes the measurements matrix and Z k denotes the measurements taken at times k-1 to k. The Gaussian measurement noise V k is represented in equation (7). N(.) with the standard probability distribution P(v).

$p(v) \sim N(O, R)$ (7)

O gives output of the system noise and R is covariance matrix for apriori estimate for measured KF.

Time update equation Equation (4) and (6) describe linear model with time k that states x_k is shown in the following equations (8) and (9).

$$x_k = Ax_{k-1} + W_{k-1} \qquad (8)$$

$$P_{\bar{k}} = AP_{k-1}A^T + Q \qquad (9)$$

The covariance error estimates Q, R with Apriori knowledge are given in the above equation (8) and (9). Convoluted moving window is applied for tracking objects in the kalman filter. For this window the covariance error estimates with Apriori knowledge is necessary. This moving filter is used to measure and process noise present in the filter to obtain the smoothened image. The noise measurement is obtained by the following equation (10)

$$\varepsilon_k = O_k - O_k^i \qquad (10)$$

Where true O_k gives true system output, and O_k gives raw measurement vector. The measure of noise ε_k estimating for O_k^i.

$$O_k^S = (1 - \omega_i)O_{k-1} + \omega_i O_k \qquad (11)$$

ω_i gives weighting factor, and smooth O_k^S gives Equation (11) smoothes time series of measurement data. Replace true O_k with smooth $O_k$in Equation (12), the estimated noise is

$$\varepsilon'_k = O_k\text{-}O_k^S \qquad (12)$$

$$E[\varepsilon'_j] = \frac{1}{N_s}\sum_{j=k-N_s+1}^{k} \varepsilon'_j \quad (13)$$

$$R_k = \frac{1}{N_s-1} \cdot \sum_{j=k-N_s+1}^{k} (\varepsilon'_j - E[\varepsilon'_j]) \cdot (\varepsilon'_j - E[\varepsilon'_j])^T \qquad (14)$$

From the above equations, the noise measured is given in equation (15) and (16)

$$\varepsilon'_k = ((1 - \omega_i).[(O_k^i - O_{k-1}^i) + (\varepsilon_k - \varepsilon_{k-1})] \qquad (15)$$

$$\varepsilon'_k \approx (1 - \omega_i).(\varepsilon_k - \varepsilon_{k-1}) \qquad (16)$$

The covariance of the actual measurement noise can be estimated in equation (17) as

$$R_k^A \approx \frac{1}{2(1-\omega_i)} R_k \quad (17)$$

The feature of the complex movable window Kalman filter is its recursive behaviour while estimating states. After the moving objects have been split, some sort of tracking procedure is necessary. First, this window must be assigned to each moving item in the scene. The picture size is somewhat greater than the complicated moving window size. It not only eliminates noise interference, but it also decreases picture processing time and improves operating speed.

These equations have to do with the system's feedback. The goal is to calculate posteriori, which would be a linear function of the priori and fresh measurement [15]. Below are the equations (18), (19), and (20).

$$C_k = P_k H^T (LP_k L^T + R_k)^{-1} \qquad (18)$$

$$\hat{x}_k = \hat{x}_{\bar{k}} + C_k(Z_k - L\hat{x}_{\bar{k})} \quad (19)$$

$$P_k = (1 - C_k L)P_{\bar{k}} \qquad (20)$$

C gives Kalman gain which is figured by over the estimation update conditions. L gives linearization network. After C_k posterior state gauge and aposterior blunder gauge P_k is processed by the estimation. The time and estimation conditions are determined recursively with past posterior gauge to foresee new aprior gauge.

CMWKF utilized for a considerable length of time recognition is characterized as far as its states, movement model and estimation condition lattice x_k is an eight layered framework state vector which can be addressed in condition (21) as

$$x_k = [x_{o,k}, y_{o,k}, l_k, h_k, v_{x,k}, v_{y,k}, v_{l,k}, v_{h,k}]^I \quad (21)$$

$x_{o,k}$and $y_{o,k}$ represent even and vertical centroid coordinate. $l_k$and h_k represent half width and half tallness of CMW. $v_{x,k}, v_{y,k}, v_{l,k}$ and $v_{h,k}$ represent their speed individually. After the state condition and estimation condition of movement model are characterized in the following casing, CMWKF will be utilized to gauge the article area

Model update When the worth of the expense work is viewed as least then the boundaries of CMWKF is should be refreshed and use them as a contribution to the following edge.

Algorithm:

To simplify the procedure, the algorithm steps are sketched and summarized below.

Step 1: Conversion of video into video frames

Step 2: Design of Gaussian Mixture Model

$$P(x_t) = \sum_{i=1}^{k}(\omega_{i,t} n(X_t; \mu_{i,t}, \sum i, t)$$

Step 3: Obtain the mixture equal mean

$$\mu_t = \sum_{i=1}^{k} \omega_{i,t}\mu_{i,t}$$

Step 4: Process Background subtraction modelling

$$B_t(x,y) = \begin{cases} 0 \; if \; |I_t(x,y) - b(x,y)| < \tau_B \\ 1f \; |I_t(x,y) - b(x,y)| \geq \tau_B \end{cases}$$

Step 5: Design of kalman filter
Step 6: Processing of kalman filter
$x_k = Ax_{k-1} + w_{k-1}$
Step 7: Estimation of Probability distribution function
$p(w) \sim N(O,Q)$
Step 8: Estimation of noise covariance
$\varepsilon_k = O_k - O_k^i$
Step 9: Design of Convoluted moving window

$$E[\varepsilon'_j] = \frac{1}{N_s} \sum_{j=k-N_s+1}^{k} \varepsilon'_j$$

Step 10: Covariance of actual noise measured

$R_k^A \approx \frac{1}{2(1-\omega_i)} R_k$

Step 11: The multiple object detected by the eight dimensional system state vector

$x_k = [x_{o,k}, y_{o,k}, l_k, h_k, v_{x,k}, v_{y,k}, v_{l,k}, v_{h,k}]^I$

4. Experimental Analysis:

In GMM CMW_KF, the tool for simulation is PHYTHON. MOT20 is the real time video dataset is considered as a dataset. The results of background suppression after processing GMM is shown in the figure 2a and 2b.

Figure 2a& 2b: The image frame taken from MOT 20-01 sequence of MOT 20 real time dataset and background suppressed image

The parameters taken for analysis are Accuracy, Precision, recall, False positive rate (FP), True positive rate (TP), Ground truth (GT), detection rate (DET), Mean Absolute position (MAP).

Accuracy Analysis:

The degree of similarity between a noise measurement and its real value is called accuracy. Table 1 represents the accuracy analysis of the proposed method

Table 1: Accuracy analysis of the proposed method

Sequence	Accuracy (%)
MOT20-01	82.86
MOT20-02	88.87
MOT20-03	84.63
MOT20-04	80.67
MOT20-05	90.69
MOT20-06	75.58
MOT20-07	89.84
MOT20-08	79.53

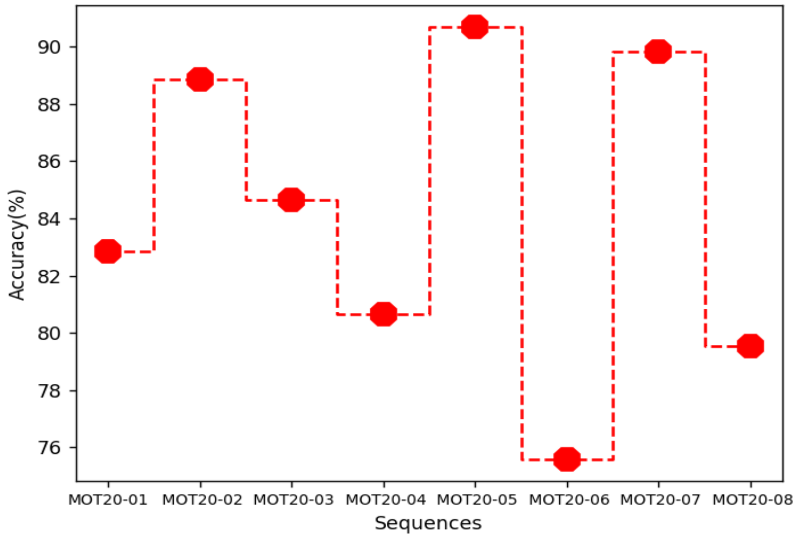

Figure 3: Accuracy analysis of the proposed method

In the figure 3 different sequences of video frames in the dataset MOT 20 are present in the X-axis and accuracy in percentage is considered in Y-axis. From this MOT20-05 achieves highest accuracy rate of 90.69%.

Precision Analysis:

The degree toward which repeated noise measurements under the same conditions get the same results is known as precision. Table 2 shows the Precision analysis of the proposed method

Table 2: Precision analysis of the proposed method

Sequence	Precision (%)
MOT20-01	99.65
MOT20-02	99.62
MOT20-03	98.72
MOT20-04	98.36
MOT20-05	99.68
MOT20-06	82.27
MOT20-07	93.78
MOT20-08	72.31

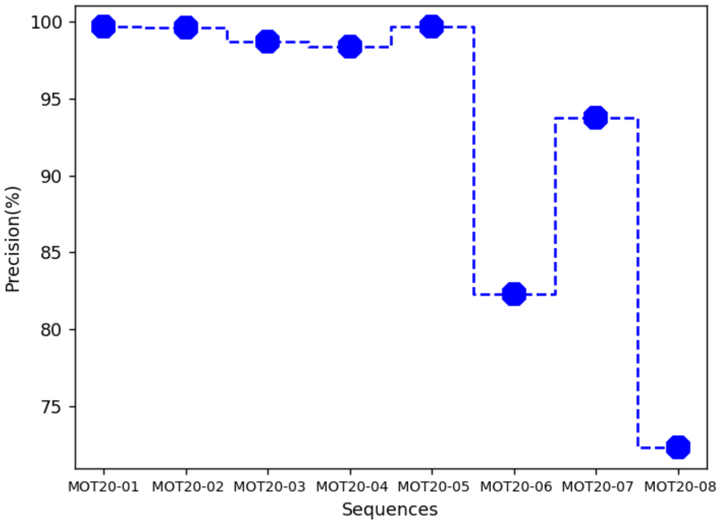

Figure 4: Precision analysis of the proposed method

In the figure 4, different sequences of video frames in the dataset MOT 20 are present in the X-axis and precision in percentage is considered in Y-axis. From this MOT20-05 achieves highest Precision rate of 99.68%.

Recall Analysis:
This is defined as the ratio of how many relevant images you have retrieved overall. Table 3indicates the Recall analysis of the proposed method

Table 3: Recall analysis of the proposed method

Sequence	**Recall (%)**
MOT20-01	89.24
MOT20-02	88.31
MOT20-03	64.3
MOT20-04	75.6
MOT20-05	70.32
MOT20-06	63.5
MOT20-07	87.8
MOT20-08	63.2

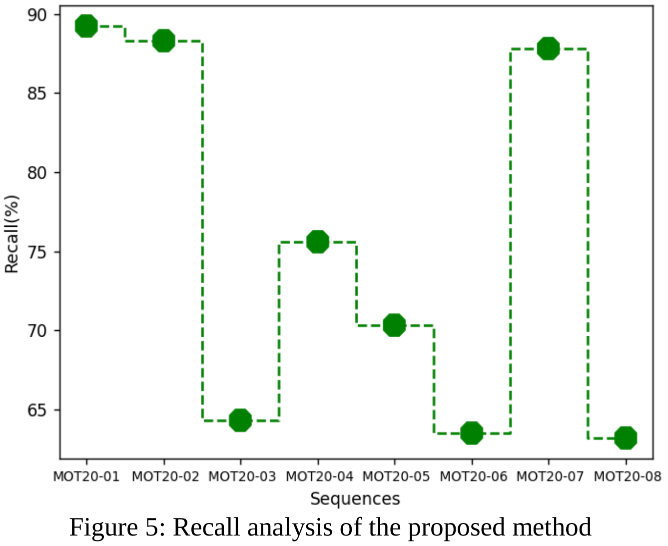

Figure 5: Recall analysis of the proposed method

In the figure 5, different sequences of video frames in the dataset MOT 20 are present in the X-axis and recall in percentage is considered in Y-axis. From this MOT20-01 achieves highest recall value of 89.24%.

True positive Analysis:

The requirements for calculating a tracker's performance. One is to see if each postulated output is a true positive (TP) that corresponds to a real (annotated) goal. Table 4 describes True positive analysis of the proposed method.

Table 4: True positive analysis of the proposed method

Sequence	**True Positive**
MOT20-01	1151
MOT20-02	9201
MOT20-03	15428
MOT20-04	11682
MOT20-05	11496
MOT20-06	3603
MOT20-07	789
MOT20-08	2059

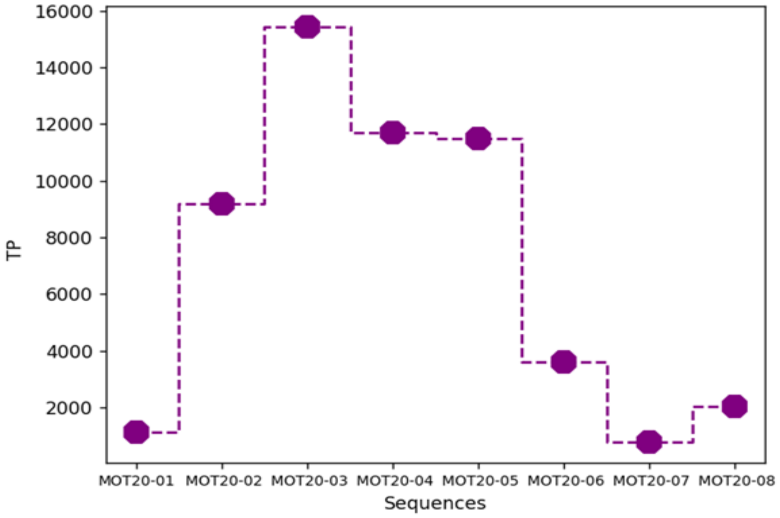

Figure 6: True Positive analysis of the proposed method

In the figure 6, different sequences of video frames in the dataset MOT 20 are present in the X-axis and true positive is considered in Y-axis. From this MOT20-03 achieves highest True Positive value of 15428.

False Positive Analysis:

The requirements for calculating a tracker's performance. The first step is to establish if each postulated outcome is just a false alarm or a false positive, FP Table 5 shows False positive analysis of the proposed method

Table 5: False positive analysis of the proposed method

Sequence	**False Positive**
MOT20-01	35383
MOT20-02	155661
MOT20-03	94209
MOT20-04	165294
MOT20-05	238859
MOT20-06	52413
MOT20-07	47148
MOT20-08	32622

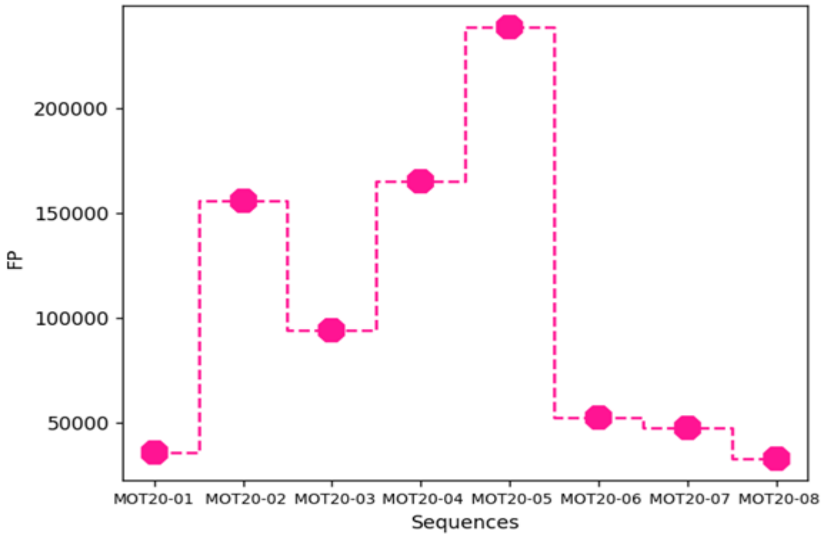

Figure 7: False Positive analysis of the proposed method

In the figure 7, different sequences of video frames in the dataset MOT 20 are present in the X-axis and false positive is considered in Y-axis. From this MOT20-05 achieves highest False Positive value of 238859.

Ground Truth Analysis:

The term "ground truth" refers to data gathered on the spot. Image data may be linked to real-world characteristics and materials using ground truth. Atmospheric adjustment is also aided by ground truth. Table 6 represents Ground Truth analysis of the proposed method.

Table 6: Ground Truth analysis of the proposed method

Sequence	**GT**
MOT20-01	12610
MOT20-02	89837
MOT20-03	177347
MOT20-04	228298
MOT20-05	381349
MOT20-06	69467
MOT20-07	20330
MOT20-08	43703

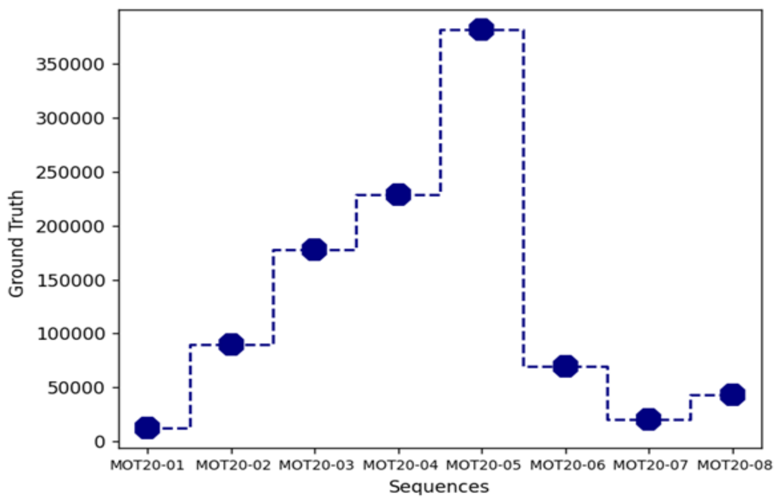

Figure 8: Ground Truth analysis of the proposed method

In the figure 8, different sequences of video frames in the dataset MOT 20 are present in the X-axis and Ground truth is considered in Y-axis. From this MOT20-05 achieves highest ground truth value of 381349.

Detection analysis:

Only pedestrians are considered in this analysis. Static people and other classes are ignored and filtered out of both the detections and the ground truth. Table 7 shows Detection analysis of the proposed method

Table 7: Detection analysis of the proposed method

Sequence	**DET**
MOT20-01	36534
MOT20-02	164862
MOT20-03	109637
MOT20-04	176976
MOT20-05	250355
MOT20-06	56016
MOT20-07	47937
MOT20-08	34681

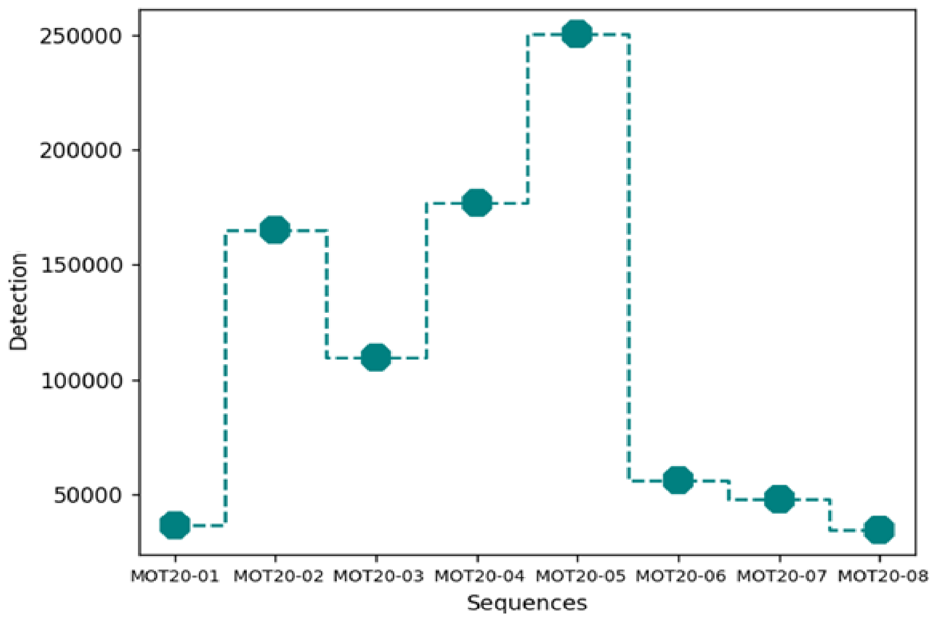

Figure 9: Detection analysis of the proposed method

In the figure 9, different sequences of video frames in the dataset MOT 20 are present in the X-axis and detection is considered in Y-axis. From this MOT20-05 achieves highest detection value of 250355.

Mean Absolute Position Analysis:

Mean Absolute Position is positional information which has details of coordinates related to a unique image. Table 7 represents Mean Absolute Position analysis of the proposed method

Table 7: Mean Absolute Position analysis of the proposed method

Sequence	MAP
MOT20-01	95.63
MOT20-02	96.3
MOT20-03	91.44
MOT20-04	92.1
MOT20-05	92.3
MOT20-06	90.3
MOT20-07	94.38
MOT20-08	87.9

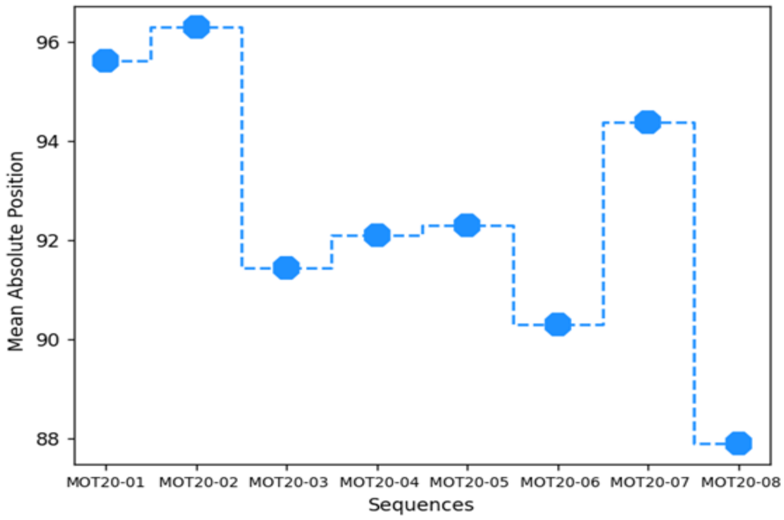

Figure 10: Mean Absolute Position analysis of the proposed method

In the figure 10, different sequences of video frames in the dataset MOT 20 are present in the X-axis and Mean Absolute Position is considered in Y-axis. From this MOT20-02 achieves highest Mean Absolute Position value of 96.3.

The experimental findings reveal that the proposed method detects and tracks foreground objects in complex and dynamic scenarios with high accuracy, robustness, and efficiency. This approach also yields noise-free smoothed pictures.

5. Conclusion

Multi object detection utilizing Gaussian Mixture Model (GMM) and background suppression has been performed. Multiple moving objects tracking using convoluted moving window Kalman filter (CMWKF) has also been performed in this paper. It can, however, cope with a variety of real-time video sequences in the MOT 20 dataset. The experimental findings reveal that the proposed method detects and tracks foreground objects in complex and dynamic scenarios with high accuracy, robustness, and efficiency. This method also produces smoothened images without noise. The obtained results are free from and it takes less computation time. There is no mistaking of the objects and if illumination changes, the tracking of moving object is not changed, this makes this system to be more efficient. This method is to be improved in future by conducting experiments in very crowded environment with more video sequences.

References

[1] Gautam, A., & Singh, S. (2019, December). Trends in video object tracking in surveillance: A survey. In *2019 Third International conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud)(I-SMAC)* (pp. 729-733). IEEE.

[2] Luo, W., Xing, J., Milan, A., Zhang, X., Liu, W., & Kim, T. K. (2021). Multiple object tracking: A literature review. *Artificial Intelligence*, *293*, 103448.

[3] Bukey, C. M., Kulkarni, S. V., & Chavan, R. A. (2017, September). Multi-object tracking using Kalman filter and particle filter. In *2017 IEEE International Conference on Power, Control, Signals and Instrumentation Engineering (ICPCSI)* (pp. 1688-1692). IEEE.

[4] Santosh, D. H., & Mohan, P. K. (2014, May). Multiple objects tracking using extended Kalman Filter, GMM and Mean Shift algorithm-A comparative study. In *2014 IEEE International Conference on Advanced Communications, Control and Computing Technologies* (pp. 1484-1488). IEEE.
[5] Farahi, F., & Yazdi, H. S. (2020). Probabilistic Kalman filter for moving object tracking. *Signal Processing: Image Communication*, *82*, 115751.
[6] Gunjal, P. R., Gunjal, B. R., Shinde, H. A., Vanam, S. M., & Aher, S. S. (2018, February). Moving object tracking using kalman filter. In *2018 International Conference On Advances in Communication and Computing Technology (ICACCT)* (pp. 544-547). IEEE.
[7] Makhmutova, A., Anikin, I. V., & Dagaeva, M. (2020, September). Object tracking method for videomonitoring in intelligent transport systems. In *2020 International Russian Automation Conference (RusAutoCon)* (pp. 535-540). IEEE.
[8] Singh, V. K., Kumar, N., & Singh, N. (2020). A hybrid approach using color spatial variance and novel object position prior for salient object detection. *Multimedia Tools and Applications*, *79*(39), 30045-30067.
[9] Mathias, A., Dhanalakshmi, S., Kumar, R., & Narayanamoorthi, R. (2021). Underwater object detection based on bi-dimensional empirical mode decomposition and Gaussian mixture model approach. *Ecological Informatics*, *66*, 101469.
[10] Ghedia, N. S., & Vithalani, C. H. (2021). Outdoor object detection for surveillance based on modified GMM and Adaptive Thresholding. *International Journal of Information Technology*, *13*(1), 185-193.
[11] Zhu, H., Zhang, G., Li, Y., & Leung, H. (2021). A novel robust Kalman filter with unknown non-stationary heavy-tailed noise. *Automatica*, *127*, 109511.
[12] Babu, P., & Parthasarathy, E. (2022). FPGA implementation of multi-dimensional Kalman filter for object tracking and motion detection. *Engineering Science and Technology, an International Journal*, *33*, 101084.
[13] Dendorfer, P., Rezatofighi, H., Milan, A., Shi, J., Cremers, D., Reid, I., ... & Leal-Taixé, L. (2020). Mot20: A benchmark for multi object tracking in crowded scenes. *arXiv preprint arXiv:2003.09003*.
[14] Dong, E., Han, B., Jian, H., Tong, J., & Wang, Z. (2020). Moving target detection based on improved Gaussian mixture model considering camera motion. *Multimedia Tools and Applications*, *79*(11), 7005-7020.
[15] Chavan, R., & Gengaje, S. R. (2017). Multiple object detection using GMM technique and tracking using Kalman filter. *Int. J. Comput. Appl*, *172*(3), 20-25.