Received: July 31, 2023. Revised: September 8, 2023.

INASS

International Journal of
Intelligent Engineering & Systems

http://www.inass.org/

Video Stream based Co-Located Pattern Recognition and Movement Tracking Based on MD-PCA and VDI-FCC Techniques

Jayaram C V[1]* **B K Raghavendra[2]**

[1]*Department of Computer Science and Engineering,*
Mysore College of Engineering and Management, Mysuru, Karnataka, 570028, India
[2]*Department of Information Science and Engineering,*
Don Bosco Institute of Technology, Bengaluru-560074, India
* Coresponding author's Email: jayaramscv77@gmail.com

Abstract: In the videos, co-location (CL) pattern recognition (PR) aids the systems administrators in effectively recognizing abandoned and detached co-located objects in the airport terminals. Nevertheless, for co-located object recognition (OR) in video streams, works are developed scarcely. Hence, this paper proposes a co-located OR with movement analysis based on mullers' decomposition-based principle component analysis (MD-PCA) and varied distance index-based farthest first clustering (VDI-FFC). Frame conversion, pre-processing, and object detection with you only look once version 8 (YOLO v8) are performed to detect the objects in the input video. Next, to determine the main and co-located objects in the frame, the features from the detected objects and reference images are extracted and matched. Afterward, to determine the pattern movement pattern, the motion estimation with rood pattern search (RPS) and pattern estimation with random walk-normalized laplacian-based hough transform (RWN-HT) is performed with the filtered objects frame. Next, the distance is estimated and with the estimated distance, the TarIf activation-based long short term memory (TI-LSTM) predicts whether the co-located object is detached or abandoned by the main object. If the object is detached/ abandoned, such objects are recognized using the MD-PCA, also, to track the location of co-located objects, movement analysis is performed with VDI-FFC. Further, the alert is sent to the administrator monitoring the video surveillance in the airport terminal based on the determined position. If a co-located object is discovered to be disconnected, the surveillance administrator is informed. The user may be informed by the administrator through a public announcement. In 6017 milliseconds, the proposed MC-PCA achieved a recognition rate of 96.75%. With VDI-FFC, the movement analysis of the associated item is completed in 334 milliseconds. In 4535ms, the RWN-HT evaluated the patterns. With 97.26% accuracy, 98.81% precision, and 0.0675 MAE, the TI-LSTM identified the availability of objects.

Keywords: Random walk-normalized laplacian-based hough transform (RWN-HT), Varied distance index-based farthest first clustering (VDI-FFC), TarIf activation-based long short term memory (TI-LSTM), Mullers' decomposition-based principle component analysis (MD-PCA), You only look once (YOLO), Co-located object recognition (COR).

1. Introduction

PR is considered as a process of classification with the advancement of technologies, as anomaly event detection [3], multi-object detection and tracking in videos are used grounded on image processing techniques. Hence, CL objects can be detected effectually with the OR and tracking technique. However, co-location PR was performed for spatial images in prevailing methodologies. Where extracting patterns from input data and categorizing them into different classes is the goal [5, 12]. The existing researchers made use of techniques of learning the unsupervised features, which have made huge enhancements in OR, for recognizing the objects. An instance might be objects, which are determined and displayed on a monitor as they pass via a verification strip at airports [6]. Since co-located

PR helps to recognize the lost objects of passengers in places, such as airports, bus terminals, railway stations, and so on, it plays a vital role in PR. The group of spatial items, which are usually found next to one another, is termed the co-located patterns [8]. The application of co-located PR helps to find the objects more effectively. For example, belongings that detach without the knowledge of individuals can be found in airport terminals. In spatial and video data, the co-located PR can be performed. Hence, OR and tracking plays a vital part to recognize the co-located object in a video.

OR, which is hugely wielded in areas like target identification and location, aircraft navigation, and space rendezvous and docking, is a significant application [9]. Object tracking works based on the detection by tracking in which the objects are detected initially and then tracking is done. In a large number of application areas [13].

Spatial CL pattern mining, which detects the feature subset whose objects are always placed together in geographic proximity, a map-reduce-centric approach [11] and an interesting as well as significant problem in spatial data mining areas [18]. The spatial CL pattern mining made use of machine learning techniques like genetic algorithm (GA) [19]. However, there is a lack of models to analyze the CL patterns in the video stream. Hence, this paper proposes a CL recognition system grounded on the MD-PCA with the VDI-FCC movement analysis in the video stream. For detecting the state of the co-located object (that is, whether the object is in static or in movement, the movement analysis is performed.

1.1 Problem statement

No works were developed based on the further issues on video stream PR as there are works developed for only spatial co-location PR. The problems are stated as,

- Lack of collection of patterns of the co-located objects, which would misclassify the other objects as co-located objects is one of the challenges of CL recognition in video streams.
- Since the video contains more objects than the main and co-located objects, the process of selection of the main object from the co-located object process was a challenging task.
- Moreover, since the co-located object would move from one place to another in crowded scenarios, the recognition of co-located objects in the video is a complex process.

By analyzing these setbacks, recognizing the co-located object in the video frame with co-located object movement analysis is the goal.

- The pattern estimation centered on the RWN-HT is proposed to rectify the lack of PR issue.
- The RANSAC feature matching based on the reference image is performed to differentiate the main and the co-located objects from the other objects.
- The movement analysis based on the VDI-FFC is proposed to determine the location of a lost co-located object.

The remaining part is arranged as: the proposed technique's related works are elucidated in section 2; a detailed explanation of the proposed techniques is delineated in section 3; the performance of the proposed algorithm is illustrated in section 4; the paper is wrapped up with the future work in section 5.

2. Related works

Baby Anitha [1] explored a hybrid model for the recognition of vehicle location, which is in motion and co-located objects. The hybrid model was presented grounded on the artificial bee colony-based neural network and dynamic threshold technique. For the hybrid model, experimentations were done, where it was detected that the system attained an enhanced accuracy level and low run time. However, the locations should be marked within the shortest path, or else the locations might be missed.

Bao [2] applied a knowledge-centric interactive post mining of user-preferred CL patterns grounded on ontologies. Also, for minimizing the number of final outputs, a pruning system was presented with a pattern filtering model. When the distance measurement for '2' co-locations was built, this technique had superior accuracy when weighed against the prevailing works owing to the consideration of composition and semantics. This technique might not detect the user-preferred patterns if the inputs encompass no-user-preferred patterns.

Chen & Sheng [4] deployed a single-object tracking centered on 2-step spatio-temporal deep features fusion in videos. A deep learning detection with kernelized correlation filtering (KCF) was deployed RANSAC by the tracking algorithm. The objects in complex scenarios were detected and tracked. Nevertheless, the techniques' performance on a small scale was worse slightly in such scenarios.

He [7] recommended a spatial CL pattern mining and pattern reconstruction system for detecting the

important spatial configurations for crime occurrence. Grounded on the spatial auto-correlation model, the CL pattern was recognized. As per the outcomes, the considerable spatial configuration was revealed with better accuracy. Yet, since the system neglected the spatial features, the spatial pattern weren't evaluated completely.

Liu [10] examined hybrid convolutional neural networks (CNNs) for small OR at airports. For incorporating small object information, a double-one path aggregation network (DO-PAN) feature fusion structure was presented. On the UAV-VIT dataset, this system attained a better mean average precision value. Nevertheless, the system wasn't applicable for the nighttime.

Rashid [14] presented a sustainable deep-learning system for OR centered on multi-layer deep features. The large-scale image and inception V3 extracted the features were detected by the very deep convolutional networks. Next, features were combined. Features were selected and then verified with the subspace discriminant analysis. A superior accuracy level was attained. However, the strong features corresponding to the object could not be extracted effectively with low-quality images.

Tran [15] employed the maximal clique and hash table (MCHT)-centric framework for the maximal prevalent CL pattern mining. For representing the neighbor relation, the maximal cliques are wielded; also, centered on the maximal cliques, the instance hash table structure was built. Hence, from the hash table, the information regarding the prevalent CL patterns was filtered. During the co-related PR, the computational time and memory consumption was minimized by the MCHT. But, the employed methodologies' reliability was limited by the unawareness of the spatial point location.

Yang [16] developed a technique for spatial co-location pattern mining with coupling relation (SCPMCR). For capturing objects inter and intra-features under the co-location restriction, the inter-coupling participating contribution index (InterPCI) intra-coupling congregation association index (IntraCAI) were developed. Moreover, for minimizing the computation time when detecting the candidate pattern, the candidate-and-search algorithm equipped with heuristic backtracking techniques were surpassed by this technique. But, the execution time was augmented while more objects were taken.

Yao [17] explored a maximal CL mining algorithm with directed road-network constraints and spatial-continuity consideration (CMDS). For augmenting the CMDS's accuracy, a network-centric prevalence index fused with distance decay effects and road direction inference was presented. When weighed against the prevailing network colocation techniques, this one was effectual and accurate. Nevertheless, the efficiency issue became a barrier when implemented in huge urban areas.

Zhu [20] presented a technique for airport small object detection centered on feature enhancement. Attentional multi-scale feature fusion enhancement (A-MSFFE) and context feature enhancement modules were presented for augmenting the image. As per the result, the small object's OR accuracy was enhanced. However, the system's temporal features were not improved.

3. Proposed methodology for the co-location object recognition in the video stream

Since the object state might be dynamic, the recognition of the missing CL object in the video stream of the airport scenario is a challenging task. Hence, to solve this problem, this paper proposes a co-located OR with movement analysis. In Fig. 1, the proposed system's block diagram is given.

3.1 Input data

Primarily, the input video of airport surveillance is taken for the co-located PR in the proposed article. Next, the input video is converted to $x$frames $(F_x)$for further processing, which is represented as,

$$F_z = \{F_1, F_2, \ldots, F_x\} \tag{1}$$

3.2 Pre-processing

Then, the pre-processing of noise removal and image enhancement using high-frequency emphasis filtering (HFEF) is performed for enhancing the frame quality. In the filtered image, the HFEF reduces the average intensity to zero. The HFEF is equated as,

$$F_z(m,n) = f^{-1}\{[\kappa_1 + \kappa_2.\hbar_{hp}(m,n)]f(m,n)\} = f^{-1}\{[\hbar_{emp}]f(m,n)\} \tag{2}$$

Here, $m,n \in F_z$ implies the pixels in the $m\ and\ n-axes$, f, f^{-1}signifies the Fourier and inverse Fourier transform, κ_1, κ_2 delineates the control parameters and $\hbar_{hp}, \hbar_{emp}$ elucidates the high pass filtered result and the emphasis filtered results, respectively. The resultant enhanced frame of HFEF is given as E_z.

3.3 Object detection

After the completion of pre-processing, object

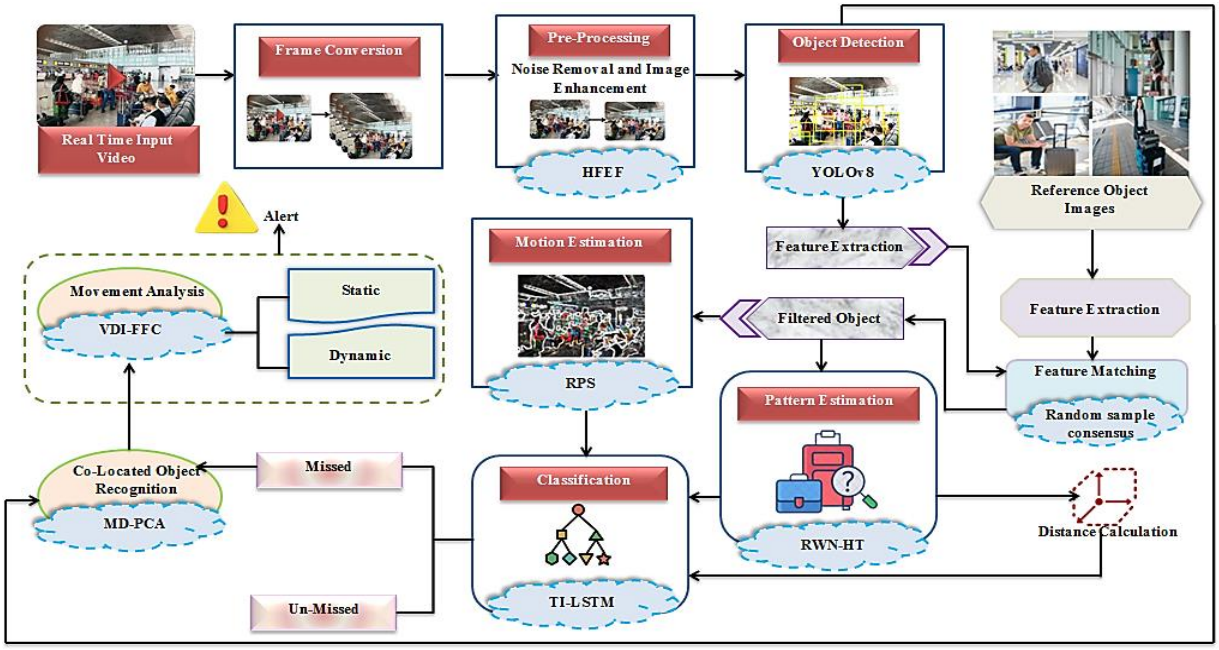

Figure. 1 Architecture of the proposed framework

detection with YOLO v8 is performed. This object detection process is to recognize the objects present in the frame and the main object and the co-located object. The main and co-located objects are referred to as the individuals and the luggage in the airport terminal.

3.3.1. Grid forming

Firstly, the enhanced frame E_Z is formed into $\varsigma \times \varsigma$grids. The anchor boxes are determined for each grid. If the center of an object in the frame falls into a grid cell, that grid is responsible for detecting that object. Hence, a bounding box is also defined for the objects in the frames E_Z .

3.3.2. Forming bounding box

The YoLo v8 predicts multiple bounding boxes per grid cell in which the head predicts the bounding box of an object and give the center coordinates (m_{center}, n_{center}) , width (ω) , and height (λ) . Hence, the final predicted bounding box (α) is estimated as,

$$\alpha_m = \sigma(A_m) + C_m \tag{3}$$

$$\alpha_n = \sigma(A_n) + C_n \tag{4}$$

$$\alpha_\omega = \beta_\omega . \exp(A_\omega) \tag{5}$$

$$\alpha_\lambda = \beta_\lambda . \exp(A_\lambda) \tag{6}$$

Here, A depicts the predicted offsets from the anchor box, C_m, C_n signifies the top-left corner of the bounding boxes coordinates, σ depicts the sigmoid activation, and $\beta_\omega, \beta_\lambda$ depicts the anchor box's height and width.

3.3.3. Overlap elimination

After the bounding boxes for an object are determined, by using Intersection of the union(I), the overlapped bounding boxes are detected.

$$I = \frac{P_\alpha \cap \Im_\alpha}{P_\alpha \cup \Im_\alpha} \tag{7}$$

Here, $P_\alpha, \Im_\alpha$ signifies the predicted and target bounding boxes. Here, the bounding box with the highest I value is selected as the final bounding box of objects. The detected objects, namely the chair, person, and luggage in the frame (δ) are expressed as,

$$\delta = \{\delta_1, \delta_2, .., \delta_v\} \tag{8}$$

Here, δ_v implies the v^{th} detected object in the frame E_Z.

3.4 Feature extraction

The reference images are used to find the main and the co-located object in the frame as the detected objects δ represent all the objects available in the scenario. The features are extracted and then matched since the reference objects cannot be directly matched with the detected objects. The scale-invariant feature transform (SIFT) and SURF

features are extracted from the object-detected frame for feature matching.

3.4.1. SIFT

Scale-space extrema detection, key point localization, orientation assignment, computation of descriptor, and key point matching are the '5' steps by which this system extracts the stable features. The scale-space image feature (S) can be obtained by,

$$S = G(d, e, \zeta) * \delta(d, e) \tag{9}$$

Here, $G(d, e, \zeta)$ signifies the Gaussian functions and $\delta(d, e)$ implies the pixel coordinates of the frame δ. Next, key points, which are sensitive to noise, are eliminated. Further, by assigning magnitude and orientation to the key points, orientation is computed. The features, which are less subjective to illumination changes, are evaluated in the computation of descriptors. In the end, to obtain the best match to the key point, Euclidean distance is measured. Hence, the determined features are depicted as $\{\wp_{ob}\}, ob = 1,2,\ldots,\Gamma$.

3.4.2. SURF

After the SIFT features are determined, the SURF feature are extracted using SURF descriptors ($\Re_s$) and is given as,

$$\Re = \delta * \chi \tag{10}$$

Here, χ implies the isotropic and separable Gaussian kernel. With this descriptor, a feature set is obtained, which is denoted as $\{\aleph_{ob}\}$. Likewise, in the reference object frame, the SURF and SIFT features are extracted and are denoted as $\{\wp_{ref}\}$, $\{\aleph_{ref}\}$. The reference image here considered is the frequent occurrences in the airport, such as persons with luggage, or backpack. From these images, features are extracted and matched with the detected object.

3.5 Feature matching

After the features are extracted, to filter the co-located and main object in the frameδ, the feature matching with random sample consensus (RANSAC) is performed. The RANSAC algorithm filters the outlier features, by this, the other objects are removed in the frame δ.

To eliminate the false matching points in features, a homography matrix (M) is deployed by the RANSAC. The relationship between the homography matrices and the feature pairs is given as,

$$\begin{bmatrix} \wp_{ref} \\ \aleph_{ref} \end{bmatrix} = \begin{bmatrix} q_0 & q_1 \\ q_2 & 1 \end{bmatrix} \cdot \begin{bmatrix} \wp' \\ \aleph' \end{bmatrix} \tag{11}$$

Here, $\wp', \aleph'$ signifies the features to be matched in $\wp_{ob}, \aleph_{ob}$, and $(q_0, q_1, q_2) \in \wp_{ob}, \aleph_{ob}$ are the parameters to be solved to obtain the filtered images. Hence, finally, a frame (Q_z) with the filtered main and co-located objects is obtained.

3.6 Motion estimation

After the filtered object (Q_Z) frame is obtained, to analyze the motion pattern of the main and the co-located objects, the motion between the frames is calculated. Here, the motion is estimated with RPS.

3.6.1. Initial search

Initially, for movement vector (MV) prediction, a set of four neighboring blocks has been considered. Here, for the primary stage, the rood arm length is equal to the predicted motion vector size. The four arms are equated as,

$$\varepsilon = \Theta\left(\sqrt{\left(\vartheta^2_{pred}(p)\right) + \vartheta^2_{pred}(y)}\right) \tag{12}$$

Here, $\left(\vartheta^2_{pred}(p)\right) + \vartheta^2_{pred}(y)$ implies the predicted motion vector's horizontal and vertical components, respectively; also, Θ signifies the round operator of the motion vector predicted.

3.6.2. Refined search

Along with the four-armed root patterns, a sixth point is added, which is the same as the predicted MV. If the current block is not the left-most block, compute the rood arm magnitude using,

$$\varepsilon = \max\left(\left|\vartheta_{pred}(p)\right|, \left|\vartheta_{pred}(q)\right|\right) \tag{13}$$

Align the rood pattern center with the search window center as well as check its four points and the position of the predicted motion vector to find the minimum error point. Repeat the search if the novel minimal error point isn't incurred at the rood pattern center else return the motion vector $\left(\vartheta_{pred}\right)$.

$$\vartheta_{pred} = \vartheta_1, \vartheta_2, .., \vartheta_b \tag{14}$$

Here, ϑ_b implies the motion vector predicted for the b^{th} frame.

3.7 Pattern estimation

After the motion is estimated in between the objects, the pattern of movement of the main and the co-located object is estimated with the RWN-HT. The hough transform (HT), which could find and follow objects even in difficult scenarios, is chosen. However, this method has the drawback of giving misleading results when the pattern lines are overlapped. Hence, to solve this problem, the random walk normalization (RWN) technique is utilized.

3.7.1. Determining hough space

A two-dimensional array with the dimensions (ρ, φ) is constructed, here φ implies the angle range, ρ signifies limited in the range $[-dl, dl]$, and dl delineates the diagonal length of the edge. Primarily, for determining the pattern of the main and the co-located object, (ρ, φ) is set to zero. The parametric equation is given as,

$$u = \cos\varphi \quad v = \sin\varphi \quad w = 0 \tag{15}$$

Here, $u, v \in \delta$ implies the u, v plane with $w \in \rho - axis$ as its normal.

3.7.2. Determining rotation matrix

Next, the rotation is applied to each point in Eq. (15) to get points for a great circle with the normal w. As the rotation takes place in one plane, a rotation matrix (L) using the RWN technique is determined as,

$$L = I - \frac{hh^T}{dl} \tag{16}$$

$$\left.\begin{aligned} h &= \frac{w-N}{\|w-N\|} \\ N &= (\cos\varphi\sin\theta \quad \sin\varphi\sin\theta \quad \cos\theta) \end{aligned}\right\} \tag{17}$$

Where, I implies the identity matrix, θ signifies the rotation angle, and h, N delineates variables used to determine the rotation matrix. Hence, the hough transformation for a point $P(\varphi, \theta) \in \delta$ is obtained by using Eqs. (16) and (17).

Hence, the final transformations (pattern) obtained (HT_{pt}) for the main (ma_{pt}) and the co-located object (co_{pt}), by estimating the rotation matrix is given as,

$$HT_{pt} = \langle co_{pt}, ma_{pt} \rangle \tag{18}$$

3.8 Distance calculation

The spatial distance between $co_{pt} and ma_{pt}$ is calculated in each frame after the pattern of the co-located object and the main object is obtained. To estimate whether the co-located object is detached from the main object, this calculation is done.

$$dist_z = \left[\left(ma_{pt}^u - co_{pt}^u \right)^2 + \left(ma_{pt}^v - co_{pt}^v \right)^2 \right] \tag{19}$$

Here, u, v implies the data in the $u, v - axes$, respectively, and $dist_z$ signifies estimated distance in the z^{th} frame. Where, the co-located object is considered as the missed object when the distance reaches maximum.

3.9 Classification

After the distance is estimated, to predict whether the co-located object is missed from the main object or not, the estimated distance $dist$ and the estimated object pattern (HT_{pt}) and the motion pattern (ϑ_{pred}) is given to the classifier. Here, the TarIf activation-based long short term memory (TI-LSTM) is leveraged for the prediction. Here, in the processing of sequential data, the LSTM is selected due to its efficiency. LSTM has a chain structure consisting of a set of subnets called memory blocks. Each block includes memory cells and three gates, such as forget, input, as well as output gates. Moreover, there is a hidden state known as short-term memory and a cell state known as long-term memory. However, due to the improper activation of the states, the LSTM has a vanishing gradient problem. The TarIf activation is used in LSTM to solve that problem. In Fig. 2, the TI-LSTM network is given.

3.9.1. Input

$dist, HT_{pt}, \vartheta_{pred}$ is collectively assigned as the input (E). The TI-LSTM receives E along with spontaneously generated two vectors within the TI-LSTM (that is, the hidden state (i_{hid}) and the cell state (i_{cel})), which is taken as the input at the time instant (j). Given the three input vectors (E, i_{hid}, i_{cel}), the TI-LSTM regulates these vectors through the gates, such as the input gate (λ_l), forgot gate (μ_l), and output gate (ϕ_l).

3.9.2. Forget gate

This gate determines whether the information from the previous cell state should be forgotten, or kept. The process of forget gate (μ_l) is given as,

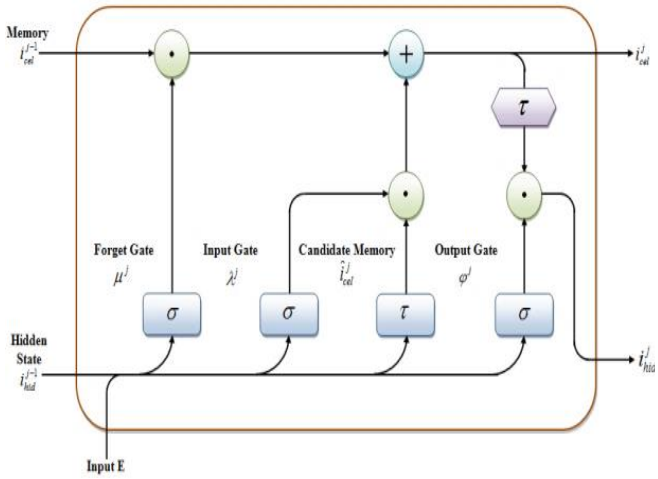

Figure. 2 Architecture of proposed TI-LSTM

$$\mu_l = \sigma\left(W_\mu\left(i_{hid}^{j-1}, E^j\right) + \varpi_\mu\right) \quad (20)$$

Here, i_{hid}^{j-1} implies the previous state input, Wsignifies the weight value, andϖdelineates the bias value.

3.9.3. Input gate

This gate is responsible to quantify the information's significance done the input gate (λ_l) and its operation is given as,

$$\lambda_l = \sigma\left(W_\lambda\left(i_{hid}^{j-1}, E\right) + \varpi_\lambda\right) \quad (21)$$

3.9.4. Output gate

The value of the successive hidden gate is determined by the output gate. The process of $\phi_l$is given as,

$$\phi_l = \sigma\left(W_\phi\left(i_{hid}^{j-1}, E\right) + \varpi_\phi\right) \quad (22)$$

3.9.5. Cell state

The state (i_{cel})aggregates the data from all the previous time steps in each memory cell of TI-LSTM and its operations are given as,

$$i_{cel} = \mu_l . i_{cel}^{j-1} + \lambda_l . \hat{\imath}_{cel} \quad (23)$$

Here, i_{cel}^{j-1} implies the previous cell state, and$\hat{\imath}_{cel}$signifies the generated new feature in the cell state, which is formulated as,

$$\hat{\imath}_{cel} = \tau\left(W_{i_{cel}} . \left[i_{hid}^{j-1}, E\right] + \varpi_{i_{cel}}\right) \quad (24)$$

Where, τimplies the TarIf activation in the TI-LSTM, which is given as,

Pseudocode of TI-LSTM

Input: Data $dist, HT_{pt}, \vartheta_{pred}$
Output: Classified co-located object availability

Begin
Initialize states (i_{hid}, i_{cel}), gates $(\lambda_l, \mu_l, \phi_l)$, time instance (j)
Set initial $i_{hid} = 0, i_{cel} = 0$
For time instant $(ti \le j)$do
Determine gates outputs $i_{cel}, \lambda_l, \mu_l$

Perform TI activation $\tau = \frac{2}{1+\exp(i_{cel})} - 1$
Update cell state with $\mu_l . i_{cel}^{j-1} + \lambda_l . \hat{\imath}_{cel}$
Estimate output gate (ϕ_l)value
End For
Return output i_{cel}^j
End

$$\tau = \frac{2}{1+\exp(i_{cel})} - 1 \quad (25)$$

3.9.6. Output

In the end, the TI-LSTM output is determined from the hidden state$\left(i_{cel}^j\right)$ as,

$$\left(i_{cel}^j\right) \qquad i_{cel}^j = \phi_l \times \tau \quad (26)$$

Here, i_{cel}^j implies the final output of TI-LSTM, which is the hidden state of the final cell. From the predicted output, it can be checked whether the co-located object is missed or not.

3.10 Object recognition

After predicting whether the co-located object is missed or un-missed, the lost object is recognized from the object detected framesδ if the co-located object is missed. MD-PCA is proposed for co-located OR. By eliminating the correlated features, the PCA is selected for its easy computation process. However, during the SVD, the PCA has the drawback of information loss. Hence, to solve this problem, Mullers' decomposition is utilized in PCA.

3.10.1. Covariance matrix construction

The MD-PCA constructs a covariance matrix to get the eigenvectors for the recognition process. The covariance matrix (R)construction is formulated as,

$$R = \frac{1}{J}\sum_{y=1}^{U}(u_y)\,(v_y)^T \quad (27)$$

Here,Uimplies the total number of objects in the frame δ, J signifies the number of co-located objects that need to be recognized, and $()^T$ depicts matrix transpose.

3.10.2. Eigen value calculation

To find the MD from the objects in the frame, the eigenvalue is calculated as,

$$B(\delta) = \Xi\left((1/J) \times u_y \times v_y\right) \quad (28)$$

Here, $B(\ )$ implies the eigenvalue, and $\Xi(\ )$ signifies the MD function, which is given as,

$$\Xi(\delta) = D_{co}D_{main} \quad (29)$$

Where, D_{co}, D_{main} implies the decomposition of two matrices of the co-located object and main object in the frame.

3.10.3. Eigen vector estimation

The eigenvector is calculated for the features with high eigenvalues. The eigenvector $(V)$is calculated using the formula,

$$V = R - \ell.B \quad (30)$$

Where,ℓimplies a random constant value.

3.10.4. Obtaining principal components

After the eigen values are estimated, the features with high eigenvalues are derived as the principal components (recognized co-located objects). The principal components are calculated using,

$$C_H = V \times k_{cen} \quad (31)$$

Here, k_{cen} implies the kernel center. Hence, the recognized objects are given as,

$$O_\gamma = [O_1, O_2, \ldots, O_v] \quad (32)$$

Here, O_v implies the v^{th} recognized co-located object.

3.11 Movement analysis

After the detached co-located objects are recognized, to determine whether the object is present in the detached place or moved from the detached place, the movement analysis for those objects is then performed. Here, the VDI-FFC is introduced for pixel-based movement analysis. Here, the FFC is considered due to its effectiveness in clustering the dynamic pixels. However, the FFC has the drawback of forming a cluster for a longer time. Hence, to solve this problem, the varied distance index (VDI) is computed.

Pseudocode of VDI-FFC

Input: Pixel points of collated objects V
Output: Static and dynamic states ν_Φ
Begin
Initialize number of pixel points V, clusters**,** number of iterations $i_{\max}$
While $(i_{\max} - 1 = i_{\max})$
Select cluster centres(ψ, ϖ)
For each V**do**
Compute inter-cluster distance η_{int}
Compute intra-cluster distance η_{inr}
Group static and dynamic pixels $[v_{static}, \nu_{dynamic}]$
End for
Return the obtained clusters ν_Φ
End While
End

Primarily, Vpixel points are selected as cluster centers in which the first center (ψ) is selected randomly and the other centers are selected as the pixel point farthest from the first center. Hence, the selected cluster centers are depicted as (ψ, ϖ).

Next, to determine the static and dynamic pixels, the distance is calculated for the remaining pixels. Here, the distance is calculated with the FFC as,

$$\eta_{\text{int}} = \sqrt{(\psi - p_{\text{int}})(\varpi - r_{\text{int}})} \quad (33)$$

$$\eta_{inr} = |\psi - r_{\text{int}}| = \left||\varpi - p_{in}|\right| \quad (34)$$

Here, $\eta_{\text{int}}, \eta_{inr}$ implies inter and intra-cluster distance, and $p_{\text{int}}, r_{\text{int}}$ signifies the pixels present within the cluster with the center $\psi\ and\ \varpi$, respectively. Until all the static and dynamic pixels are grouped separately, this distance computation is performed. Thus, the final cluster obtained is given as,

$$\nu_\Phi = [v_{static}, \nu_{dynamic}] \quad (35)$$

Where, $v_{static}, \nu_{dynamic}$ implies the co-located

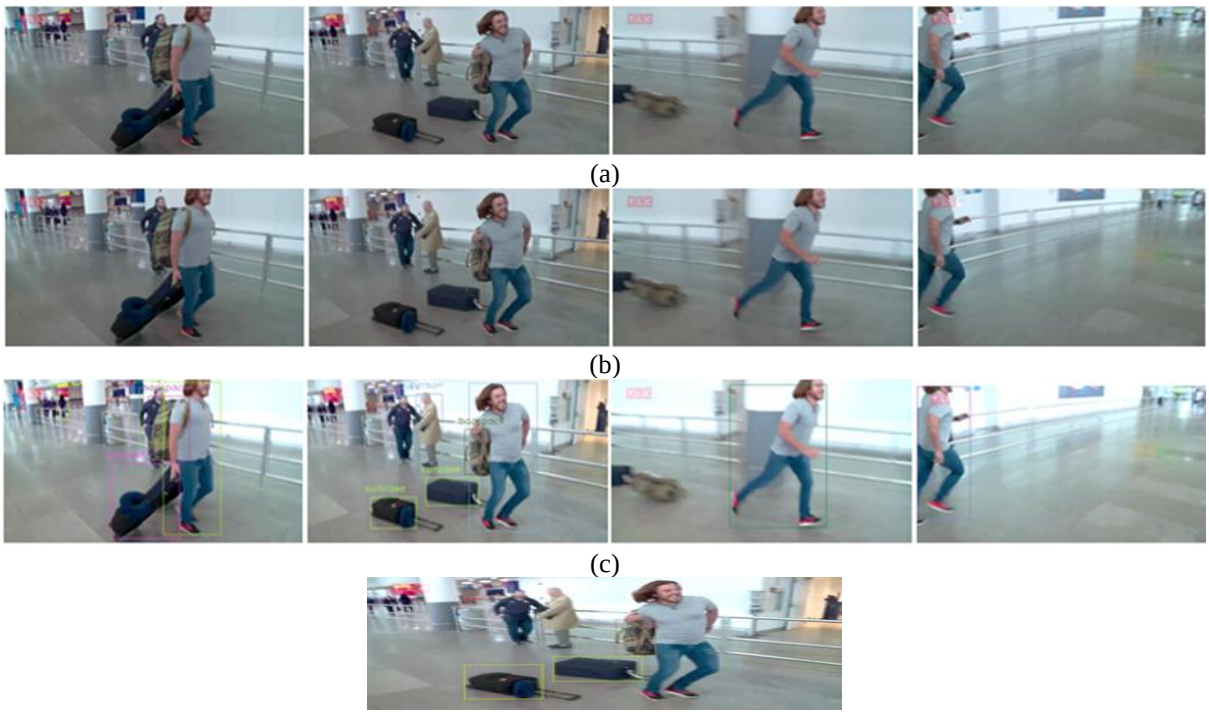

Figure. 3 Sample images of: (a) Converted frames, (b) Contrast-enhanced images, (c) Object detected images, and (d) recognised co-located objects

object in a static state and dynamic state, respectively. In the end, further, the performance of the proposed approaches is experimentally analyzed.

After determining the static or dynamic location of the co-located object, regarding location of the co-located object, the alert is sent to the administrator monitoring the airport terminal surveillance.

4. Results and discussions

Here, the proposed techniques' performance is experimentally analysed to prove their reliability. The experiments were performed on the working platform of PYTHON with the dataset collected from common object in context (COCO) available sources. In Fig. 3, the sample image results of the proposed COR in the airport scenario are given.

4.1 Performance analysis of movement estimation

Here, the proposed VDI-FFC's performance is evaluated with the prevailing FFC, K-nearest neighbor (KNN), balanced iterative reducing and clustering using hierarchies (BIRCH), and clusterin LARge applications (CLARA).

The time taken by the proposed VDI-FFC and existing FFC, KNN, BIRCH, and CLARA methods for movement analysis is given in Fig. 4. The proposed method reported 263 ms faster than the existing FFC method in the assessment of co-located object motion. Thus, VDI-based clustering can able

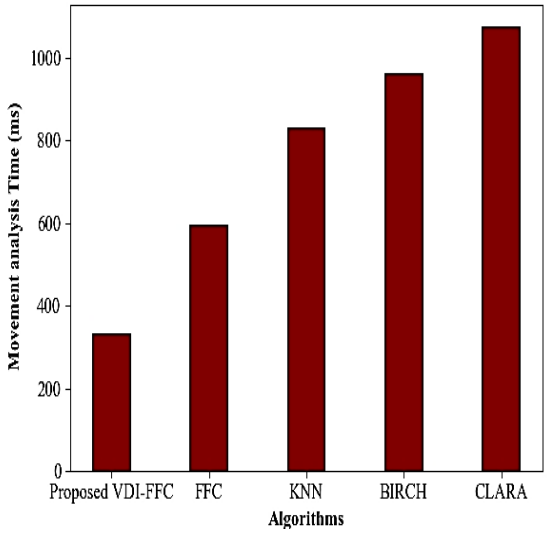

Figure. 4 Performance comparison for movement analysis

to provide rapid and precise results for identifying and measuring the movement of the co-located objects. So, the proposed method examines the movement of objects more qualitatively and quickly than the existing methods.

4.2 Performance analysis of co-located object recognition

Here, regarding recognition rate and recognition time, the performance co-located OR by the proposed MD-PCA is analyzed in comparison with the

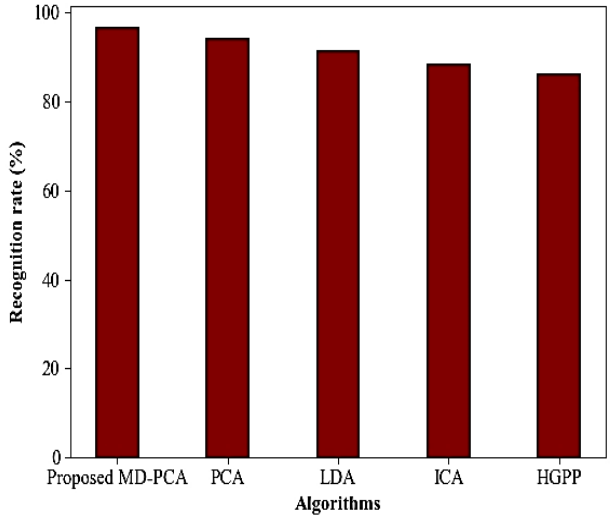

Figure. 5 Analysis of recognition rate

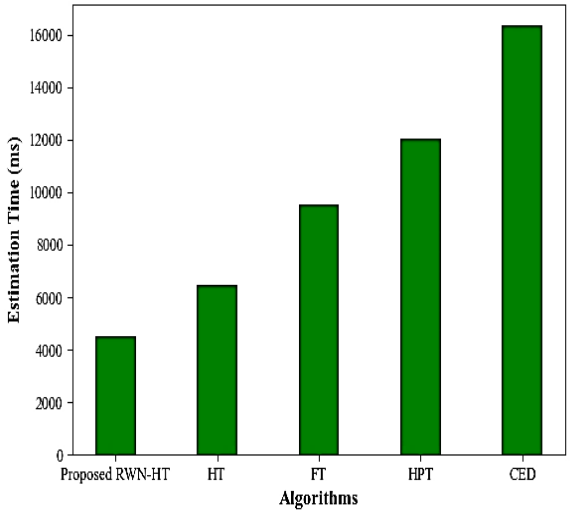

Figure. 6 Performance analysis based on estimation time

Table 1. Experimental outcome of recognition time

Techniques	Recognition time (ms)
Proposed MD-PCA	6017
PCA	10129
LDA	19012
ICA	23894
HGPP	30653

Table 2. RMSE analysis

Algorithms	RMSE
Proposed RWN-HT	1.96
HT	2.73
FT	3.48
HPT	3.99
CED	4.51

prevailing PCA, linear discriminant analysis (LDA), independent component analysis (IDA), and histogram of gabor phase patterns (HGPP).

In Fig. 5, the recognition rate attained by various methods is shown. The proposed model outperformed the existing methods with a maximum value of recognition rate i.e., 96.75%. But, the existing methods attained a lesser value of recognition rate in comparison. Thus, the OR by the proposed method is more accurate than the existing methods with the accurate classification of the availability of the co-located object.

In Table 1, the performance analysis of the proposed MD-PCA and the prevailing methods are represented regarding the recognition time of the network. The time consumption should be low to prove the model's efficacy. The proposed MD-PCA utilized a recognition time of 6017ms, while the prevailing models exhibited an increased time of 4112 ms for PCA, 12995ms for LDA, 23894ms for ICA, and 24636ms for HGPP. Thus, the proposed model has recognized the co-located object with minimum time.

4.3 Performance analysis of pattern estimation

Here, regarding estimation time and root mean square error (RMSE), the proposed pattern estimation algorithm RWN-HT's performance is analysed with the conventional HT, fourier transform (FT), hexagonal pattern transform (HPT), and canny edge detection (CED).

In Fig. 6, it is unveiled that to estimate the patterns, the HT technique takes 6487ms, which is lower than the FT, HPT, and CED techniques. Nevertheless, the estimation time is lowered by 1952ms with the RWN technique in the HT. This proves that the main and co-located object patterns can be detected in less time with the proposed model.

The RMSE outcome of the proposed and the traditional pattern estimation algorithms are illustrated in Table 2. The RMSE value of the proposed RWN-HT is 1.96, which is the least error value obtained when compared to the conventional HT, FT, HPT, and CED techniques. Hence, the RWN-HT can estimate the pattern with less error.

4.4 Performance analysis of classification

Here, regarding accuracy, precision, recall, f-measure, specificity, mean absolute error (MAE), false negative rate (FNR), false discovery rate (FDR), false positive rate (FPR) and area under the curve (AUC), the proposed TI-LSTM's performance is evaluated with LSTM, recurrent neural network

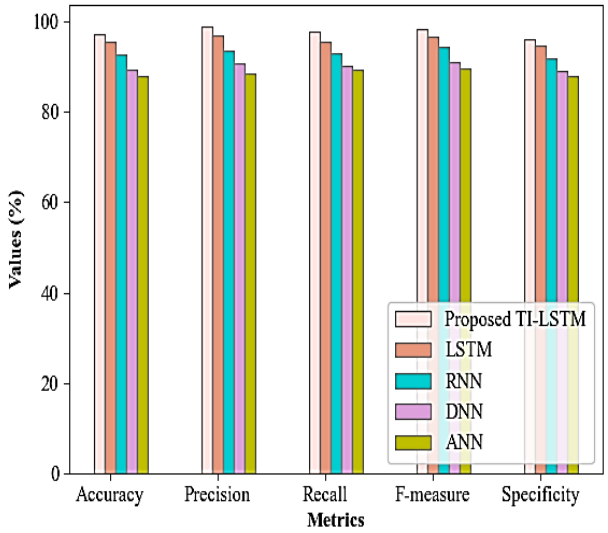

Figure. 7 Performance measure of proposed TI-LSTM

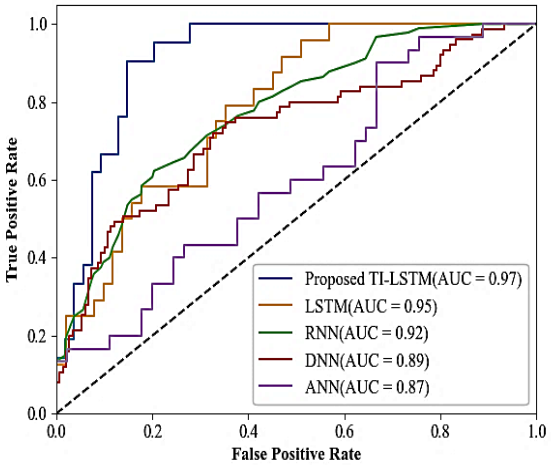

Figure. 8 AUC analysis

(RNN), deep neural network (DNN), and artificial neural network (ANN).

In Fig. 7, regarding accuracy, precision, f-measure, recall, and specificity, the proposed TI-LSTM's performance is evaluated with the prevailing techniques. The proposed TI-LSTM attains an accuracy of 97.26%, whereas the existing models attain an accuracy of 95.42% for LSTM, 92.75% for RNN, 89.22% for DNN, and 87.94% for the ANN, which are lower compared to the proposed model. Likewise, the precision, F-measure, recall, and specificity of the proposed model exhibit a superior performance of 98.81%, 98.22%, 97.65%, and 96.08%, respectively. Thus, the inclusion of Tarif in a conventional LSTM network has enhanced the accuracy of the predicted classes.

The efficiency of the proposed TI-LSTM in terms of the AUC is highlighted in Fig. 8. The AUC attained by the proposed TI-LSTM is better (0.97%) when compared with the conventional LSTM

Table 3. Experimental results of proposed and existing classifiers based on MAE

Techniques	MAE
Proposed TI-LSTM	0.0675
LSTM	0.0937
RNN	0.1185
DNN	0.1267
ANN	0.1462

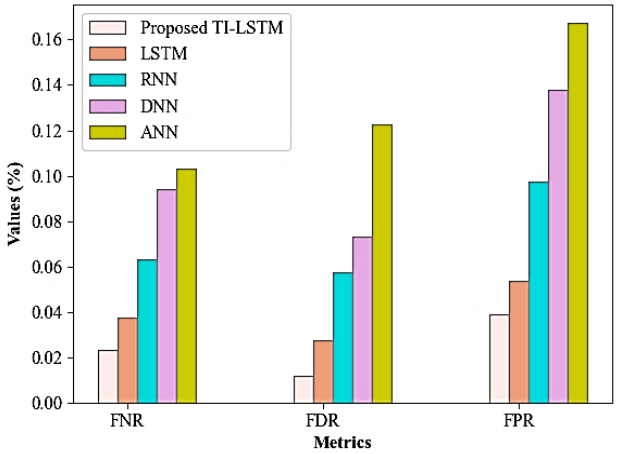

Figure. 9 Performance evaluation based on FNR, FDR, and FPR

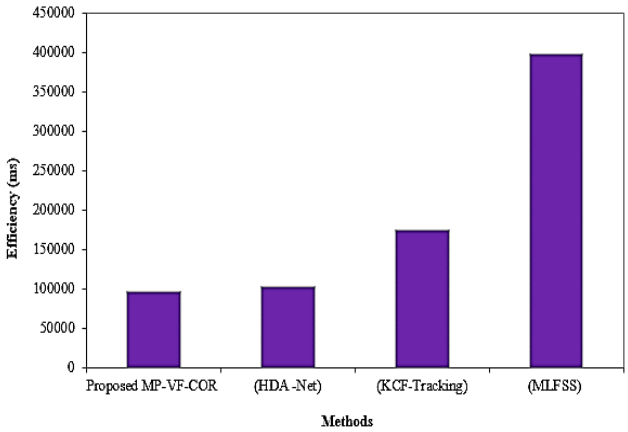

Figure. 10 Comparative analysis with existing models

(0.95%). Similarly, the AUC attained by the rest of the techniques like RNN (0.92%), DNN (0.89%), and ANN (0.87%) is also lower than the proposed method. Hence, the proposed TI-LSTM provides efficient location recognition with better performance metrics.

The MAE of proposed TI-LSTM and existing methods is depicted in Table 3. The absolute average error between the real data and the predicted data are measured by the MAE, which should be lower for better performance of the system. The MAE of the TI-LSTM is lesser by 0.0262 than the existing LSTM. Likewise, the TI-LSTM has performed better. Thus, the TarIf activation in the LSTM technique reduced

the likelihood of vanishing gradient problems, which provides better results.

In Fig. 9, the error values of the proposed TI-LSTM are displayed in terms of FPR, FNR, and FDR. The lower value of FDR, FPR, and FNR, depicts the better prediction rate of this system. The proposed method achieves 0.0119 of FDR, 0.0392 of FPR, and 0.0235 of FNR values, whereas the average FPR, FDR, and FNR values attained by the existing DNN technique are 0.0734, 0.139, and 0.0942, respectively. Thus, the proposed algorithm predicted wrongly in only less number of cases.

4.5 Comparative analysis

Here, the proposed MD-PCA-VDI-FFC –based COR (MP-VF-COR) evaluated with the existing works of KCF-kernelized correlation filtering model [4], hybrid convolutional structure, double-one PAN [10] and AIoU loss (HDA-Net) and multi-layers deep features fusion and selection regarding efficiency [14].

The efficiency is analyzed in terms of time for the recognition process is depicted in Fig. 10. To track the objects, [4] KCF-kernelized correlation filtering model takes 7814ms higher than the proposed model. Whereas [14] for detecting the objects only, the multi-layers deep features fusion and selection regarding efficiency and [10] hybrid convolutional structure, double-one PAN, and AIoU loss (HDA-Net) models take 303358ms and 6214ms time. Hence, the proposed model is more efficient for co-located OR with movement analysis.

5. Conclusion

This work performs the co-located object PR in the video frame with the movement tracing. TI-LSTM and RWN-HT are proposed for classification and pattern estimation. MD-PCA and VDA-FFC were deployed for the co-located OR and movement tracking. An alert message is sent to administer who supervises the surveillance of the airport terminal regarding the detached or abandoned co-located object. In the end, the experiments were conducted on the collected dataset from publically available sources. As per the experiments, the proposed MC-PCA attained a recognition rate of 96.75% in 6017ms. The movement analysis of the co-located object is performed with VDI-FFC in 334ms. The RWN-HT estimated the patterns in 4535ms. The TI-LSTM classified the availability of objects with 97.26% accuracy, 98.81% precision, and 0.0675 MAE. In the end, the time efficiency of the proposed model is proved by the comparative analysis with the existing work. Here, the recognition of co-located objects is based on movement analysis and not focused on the occlusion problem. Hence, the occlusion problems will be focused in the future and the proposed COR model.

Notation table

Parameters	*Representation*
F,x	Input video and converted frames
$m, n \in F_z$	Pixel in m and n axes
f, f^{-1}	Fourier and inverse transforms
$\beta_\omega, \beta_\lambda$	Anchor box's height and width
σ	Sigmoid Activation
$P_\alpha, \Im_\alpha$	Predicted and target bounding boxes.
(S)	The scale-space image feature
$G(d, e, \zeta)$	Signifies the Gaussian functions
χ	Implies the isotropic and separable Gaussian kernel.
φ	Angle range,
θ	Rotation angle
(ϑ_{pred})	Motion pattern
$(i_{hid}), (i_{cel})$	Hidden state and the cell state
$(\lambda_l), (\mu_l), (\phi_l)$	Input gate, forgot gate and output gate.
(R)	Covariance matrix
U	Total number of objects in the frame δ
J	Number of co-located objects
$B(\ \ ), \Xi(\ \ )$	Eigen value and MD function,
k_{cen}	Kernel center
O_v	v^{th} recognized co-located object.
$\eta_{\text{int}}, \eta_{inr}$	Inter and intra-cluster distance
$v_{static}, v_{dynamic}$	Co-located object in a static state and dynamic state,

Conflicts of interest

The authors declare no conflict of interest.

Author contributions

The paper background work, conceptualization, methodology, dataset collection, implementation, result analysis and comparison, preparing and editing draft, visualization have been done by first author. The supervision, review of work has been done by second author.

References

[1] E. B. Anitha, S. Sivaprakash, S. Velmurugan, and S. S. Saranya, "Hybrid artificial bee colony based neural network and dynamic threshold technique for predicting moving vehicle location

and co-located objects", *Sadhana-Academy Proceedings in Engineering Sciences,* Vol. 48, No. 2, pp. 1-9, 2023.

[2] X. Bao, T. Gu, L. Chang, Z. Xu, and L. Li, "Knowledge-based interactive postmining of user-preferred co-location patterns using ontologies", *IEEE Transactions on Cybernetics,* Vol. 52, No. 9, pp. 9467–9480, 2022.

[3] Y. Chang, Z. Tu, W. Xie, B. Luo, S. Zhang, H. Sui, and J. Yuan, "Video anomaly detection with spatio-temporal dissociation", *Pattern Recognition,* Vol. 122, No. 108213, 2022.

[4] Y. Chen and R. Sheng, "Single-object tracking algorithm based on two-step spatiotemporal deep feature fusion in a complex surveillance scenario", *Mathematical Problems in Engineering,* pp. 1-11, 2021.

[5] P. F. Viger, W. Gan, Y. Wu, M. Nouioua, W. Song, T. Truong, and H. Duong, "Patternmining: current challenges and opportunities", *Lecture Notes in Computer Science,* pp. 34–49, 2022.

[6] D. I. Gota, A. Puscasiu, A. Fanca, H. Valean, and L. Miclea, "Threat objects detection in airport using machine learning", In: *Proc of the Proceedings of the 21st International Carpathian Control Conference,ICCC* ,2020.

[7] Z. He, M. Deng, Z. Xie, L. Wu, Z. Chen, and T. Pei, "Discovering the joint influence of urban facilities on crime occurrence using spatial co-location pattern mining", *Cities,* Vol. 99, No. 102612, 2020.

[8] C. V. Jayaram and N. U. Bhajantri, "A review study: recognition of co-located pattern in video stream", In: *Proc. of 5thInternational Conference on Advanced Computing and Communication Systems*, pp. 353–356, 2019.

[9] Q. Jiang, Y. Yang, M. Zhang, and S. Zhang, "A line objects recognition algorithm based on non-subsampled contourlet transform", *Chinese Control Conference*, pp. 3050–3055, 2020.

[10] S. Liu, R. Wu, J. Qu, and Y. Li, "HDA-Net: Hybrid convolutional neural networks for small objects recognization at airports", *IEEE Transactions Instrumentation and Measurement*, Vol. 71, pp. 1–14, 2022.

[11] S. Maiti and R. B. V. Subramanyam, "Mining co-location patterns from distributed spatial data", *Journal of King Saud University - Computer and Information Sciences,* Vol. 33, No. 9, pp. 1064–1073, 2021.

[12] M. Paolanti and E. Frontoni, "Multidisciplinary pattern recognition applications:A review", *Computer Science Review*, Vol. 37, No. 100276, 2020.

[13] A. Pramanik, S. K. Pal, J. Maiti, and P. Mitra, "Granulated RCNN and multi-class deep SORT for multi-object detection and tracking", *IEEE Transactions on Emerging Topics in Computational Intelligence*, Vol. 6, No. 1, pp. 171–181, 2022.

[14] M. Rashid, M. A. Khan, M. Alhaisoni, S. H. Wang, S. R. Naqvi, A. Rehman, and T. Saba, "A sustainable deep learning framework for object recognition using multi-layers deep features fusion and selection", *Sustainability,* Vol. 12, No. 12, pp.1-22, 2020.

[15] V. Tran, L. Wang, H. Chen, and Q. Xiao, "MCHT: A maximal clique and hash table-based maximal prevalent co-location pattern mining algorithm", *Expert Systems with Applications*, Vol. 175, No. 114830, 2021.

[16] P. Yang, L. Wang, X. Wang, and L. Zhou, "SCPM-CR: A novel method for spatial co-location pattern mining with coupling relation consideration", *IEEE Transactions on Knowledge and Data Engineering,* Vol. 34, No. 12, pp. 5979–5992, 2022.

[17] X. Yao, X. Jiang, D. Wang, L. Yang, L. Peng, and T. Chi, "Efficiently mining maximal co-locations in a spatial continuous field under directed road networks", *Information Sciences*, Vol. 542, pp. 357–379, 2021.

[18] J. S. Yoo, D. Boulware, and D. Kimmey, "Parallel co-location mining with MAPREduce and NoSQL systems", *Knowledge and Information Systems*, Vol. 62, No. 4, pp. 1433–1463, 2020.

[19] L. Zhao and R. Jun-yi, "Incremental mining method of spatial co-location pattern based on genetic algorithm", In: *Proc. of International Conference on Industrial Application of Artificial Intelligence*, pp. 254–258, 2020.

[20] X. Zhu, B. Liang, D. Fu, G. Huang, F. Yang, and W. Li, "Airport small object detection based on feature enhancement", *IET Image Processing*, Vol. 16, No. 11, pp. 2863–2874, 2022.